CUK-H06054-18-P1504.

THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

⑱

FEBRUARY 1977

VOL. 47, PART 1

Vol. 47

THE
BRITISH JOURNAL
OF
EDUCATIONAL PSYCHOLOGY

Edited by
N. J. ENTWISTLE

Assistant Editors :
S. N. BENNETT
G. BROWN

Sub-Editor :
S. M. CUNNINGHAM

Advised by

L. B. BIRCH
E. A. PEEL
P. E. VERNON

Vol. 47, Part 1

The British Journal of Educational Psychology is issued by the British Psychological Society once a term, in February, June and November. The subscription is £8·00 a year post free, payable in advance, or £3·00 each number.

Subscriptions and Orders should be sent to Scottish Academic Press, 33 Montgomery Street, Edinburgh, EH7 5JX. Members of the British Psychological Society receive the *Journal* on special terms.

Papers for publication should be sent to Professor N. J. Entwistle, Department of Educational Research, University of Lancaster, Lancaster. Guidance to contributors on the preparation of papers for the *Journal* is given at the end of the November number each year.

All *Books for Review* for the *British Journal of Educational Psychology* should be sent to Professor N. J. Entwistle at the above address. *Advertisements* for this *Journal* should be sent to Scottish Academic Press.

Published by
SCOTTISH ACADEMIC PRESS, 33 MONTGOMERY STREET, EDINBURGH, EH7 5JX,
for the British Journal of Educational Psychology, Ltd.

All rights reserved.

© 1977. The British Journal of Educational Psychology.

PRINTED BY BIRMINGHAM PRINTERS, LTD., 8-24 SAMPSON ROAD NORTH, BIRMINGHAM.

370.1505
B 777

EDITORIAL

OWNERSHIP OF THE JOURNAL

The *Journal* has been controlled since its formation in 1931 by an independent company, British Journal of Educational Psychology Limited. Shares in the company were equally divided between the British Psychological Society and an organisation which subsequently became the Association of Teachers in Colleges and Departments of Education. The division of ownership was intended to ensure a full representation of education and teacher-training interests. Before its closure in December, 1975, the ATCDE agreed to transfer its shareholdings to the BPS on the understanding that its 'constituency' would continue to be represented on the Editorial Board of the *Journal*. The Council of the BPS has agreed in principle to an arrangement under which the *Journal* would be published by the Society, but guaranteeing a substantial degree of control of its own finances to the Editorial Board for a period of seven years. It has also agreed to continued representation of educational and teacher-training interests on the Editorial Board. Thus while, in due course, it may prove necessary to wind up the company, this will in no way affect the publication of the *Journal*. It is hoped that continuing editorial and financial independence will have been assured under the new arrangements.

18

Br. J. educ. Psychol., 47, 1-6, 1977

SYMPOSIUM :

PRACTICAL AND THEORETICAL ISSUES IN PIAGETIAN PSYCHOLOGY—III

PIAGET'S PSYCHOLOGY IN PRACTICE*

By J. SMEDSLUND
(*Institute of Psychology, University of Oslo, Norway*)

SUMMARY. As a result of eight years' experience of Piagetian research and subsequent work as a practising psychologist, certain aspects of Piaget's theories are criticised. In particular the existential status of operatory structures, the relation between logicality and understanding, and the representativeness of the tasks are questioned. Piagetian theory is seen as forcing the psychologist to adopt an artificially detached and one-sidedly cognitive attitude towards children, and to focus interest entirely on very abstract aspects of performance. While accepting the historical importance of Piaget's contributions, it is argued that the limitations and constraints implicit in the theory seriously threaten its usefulness in practical situations.

INTRODUCTION

IN this paper I will regard the psychology of Piaget not primarily as it is embodied in books or articles, but as leading to a form of practice involving mainly certain types of diagnostic encounter with children. In other words, I will be primarily concerned with what Piagetian psychologists *do*, rather than with how they write and talk *about* things. Starting from a consideration of this practice, I will then proceed to draw certain theoretical conclusions.

My first encounter with Piagetian psychology in practice occurred in the attic of École International in Geneva in the autumn of 1957. I still remember vividly the excitement of being free to *converse* with my subjects and particularly being free to ask for their *explanations*. As a well trained experimental psychologist I had up to that time been mainly confined to marks on sheets of paper as my data. Given my empiricist background I was also much intrigued and impressed by the children's apparent belief in the *necessity* of certain relations, and in their later dealings with what is merely *possible*. Altogether, I felt somehow liberated from a metatheoretical straitjacket. Some, but not all, of the discrepancy between what I felt to be legitimate research interests and the phenomena of my everyday life were removed.

Including my year in Geneva, I spent altogether about eight years doing research within the Piagetian framework. During that time I visited and worked in between 50 and 70 schools, kindergartens and nursery schools in Switzerland, USA, and Norway. More than 1,400 children participated in my published experiments and I saw hundreds of others during informal explorations and pilot studies.

Very gradually, I became aware of what I think are some major weaknesses inherent in this research tradition, which restrict its usefulness for practical purposes. As this became clear to me, I left the Piagetian theoretical framework and moved on towards other modes of understanding. I take this occasion to try to summarise very briefly both my main criticisms of Piaget's psychology as a type of practice, and those aspects of this practice which I still regard as valuable.

---

* This article was prepared for the special issue of the *Bulletin de Psychologie*, honouring Piaget's 80th birthday.

CRITIQUE OF PIAGETIAN PRACTICE

In my work with children I eventually came to wonder about three aspects of our encounters, namely the existential status of Piagetian operatory structures, the relation between logicality and understanding, and the representativeness of the tasks. In other words, I became concerned about what I now see as respectively some ontological, epistemological and pragmatic aspects of Piaget's position.

The existential status of Piagetian operatory structures.

My conversations and dealings with children never quite convinced me that their behaviour could be adequately described as reflecting the presence or absence of certain operatory structures. Partly, this was a matter of intuitive impression. I found that I could not deal with the children as if they had such operatory structures in the same natural and confident way that I dealt with them as being conscious, as perceiving, thinking and feeling, and utilising the words of everyday language. The preoccupation with structure and the distinction between form and content alienated me from my subjects, just as did the artificiality of the tasks, to be discussed later.

As I accumulated experimental data and read the publications of others, my doubts further increased. The empirical evidence did not provide much direct support for the existence of operatory structures. All kinds of discrepancies crop up with children of all ages and with adults, and with all kinds of concepts and structures. A child behaves in one way in one situation and in another way in another situation which may appear strictly equivalent to the first situation as far as task structure is concerned. In general, logical task structure does not seem to be a good predictor of behaviour across situational variations. This does not apply only to children but to adults as well. (See, for example, Wason and Johnson-Laird, 1972, and Wason, 1976.)

In Piaget's system the concept of horizontal décalage and resistance of the situation is invoked to explain such discrepancies. This concept has a relatively feeble explanatory power, since there are few generally reliable indicators of degree of resistance of new situations to an alleged operatory structure. Nevertheless, Piaget seems to hold some hope for future developments in this area (Piaget, 1962).

However, my objections to Piaget's logical formalism had deeper roots than the sense of alienation from the children that it brought me, and the weak and inconclusive experimental support it could muster. I became increasingly uneasy with the conception of operations as internalised versions of certain very general sensory-motor acts. The act of uniting or bringing together is such an act as well as its opposite, the act of separation or bringing apart. The assumption of such very general actions seems to reflect an *a priori* ontological position of *conceptual realism*. There are two main symptoms of this in Piaget's writings. The first is his treatment of resistance of the object or situation as a real psychological factor. One can only speak of resistance in this context if one presupposes the existence of that which is resisted, namely, the operatory structure. The other symptom is the terminology in which the operative structures are described. There are said to exist only a few of these structures, such as the additive and multiplicative ' groupings ' of classes and relations. Each of one of them, such as *the* additive ' grouping ' of classes, is supposed to be activated in all those infinitely numerous situations which have the corresponding structure, but nothing else, in common. If the child does not behave in accordance with the structure, this is explained by assuming too strong resistance from the situation.

Piaget's assumption of the existence of such purely structural entities as the additive grouping of classes, etc., is unacceptable to me, because it implies that human beings can function abstractly, i.e., out of context and independently of content. Everyday experience, as well as numerous experiments, show that this is not the case.

The alternative to Piaget would be to regard only children's concrete activities in concrete situations as real. This would involve no distinction between form and content and would eliminate the need for a concept of resistance, since there would be nothing to resist. The explanation of why the child behaves in a given way in a given situation would have to be *historical.* An example of this was a little boy in one of my pilot studies who failed in most of the tasks, but was superb in the task of measuring length. The boy's father turned out to be a carpenter, and the boy had been allowed to watch and help his father many times.

Piaget's theory characterises children of various ages on the basis of purely structural criteria. What generalised descriptions of children can be produced if one takes an alternative position? The answer is that one can predict the sequence of many acquisitions (*a*) from knowledge of the organisation of the various tasks (some skills and concepts are presupposed by other skills and concepts, etc.), and (*b*) from knowledge of the institutionalised sequences of instruction in a given culture. Prediction of characteristics of children of various ages across cultures has to be very uncertain, and rely only on whatever general similarities in experiences there may be and perhaps on some general characteristics of maturation linked with the species.

In summary, I gradually came to regard Piaget's abstract formalism as counterintuitive, as somewhat alienating in practice, and as without sufficient scientific foundation.

The relation between logicality and understanding.

During the prolonged debates about criteria for the presence or absence of certain structures, notably conservation and transitivity (for references, see Smedslund, 1969), I came to recognise a problem which seems to have no satisfactory solution within Piagetian psychology. In order to decide whether a child is behaving logically or not, one must take for granted that he has correctly understood all instructions and terms involved. On the other hand, in order to decide whether or not a child has correctly understood a given term or instruction, one must take for granted that the child is behaving logically with respect to the implications which constitute his understanding. The correct understanding of a given message is checked by observing agreement or disagreement with respect to what is equivalent to, implied by, contradicted by or irrelevant to the message (Smedslund, 1970). Hence, one must presuppose logicality in such respects. There is a circular relation between logicality and understanding, each one presupposing the other, and this constraint forces the researcher to make a choice of which one to take for granted and which one to study.

I gradually came to realise that the only defensible position is always to presuppose logicality in the other person and always to treat his understanding of given situations as a matter for empirical study. From this point of view, people are always seen as logical (rational) given their own premises, and hence behaviour can, in principle, always be understood. This also applies to small children, foreigners and psychotics. The opposite position, namely, always taking understanding for granted and always treating logicality as a matter of empirical study, is taken by no one. However, an intermediate inconsistent position seems to be characteristic of much of academic psychology including the Piagetian variant. Sometimes understanding seems to be the focus of study and logicality is taken for granted and sometimes the opposite seems to be the case. Only when presence or absence of logicality is studied and understanding is taken for granted, does it become legitimate to describe children as ' non-conservers,' adults as ' mad,' foreigners as ' inconsistent and self-contradictory,' and so on. However, it is seldom recognised clearly that such decisions about logicality presuppose that one understands properly and fully the meanings of the subject's various expressions, or that checks on such understanding again presupposes logicality on the part of the subject.

It is possible and intuitively plausible always to presuppose logicality, whereas it is clearly absurd always to presuppose understanding. Intermediate positions are untenable because they have no rules for when to presuppose understanding and when to presuppose logicality, and furthermore, they always lead to infinite regressions with no rules for when to stop.

As a result of this reasoning, when I meet a small child I always take for granted that, within his limited sphere of activity and given his own premises, he is logical, and my problem is to understand what his expressions mean, and hence to grasp his existential situation. In so far as Piagetian psychologists focus on logicality as a variable (e.g., conserver or non-conserver) and give only peripheral attention to the problem of determining children's understanding of instructions and situations, I think they are making an epistemological error and are out of step with everyday human life as well as with all useful psychological practice. It may be objected that Piaget is not really denying the logicality of children at any stage, but is merely studying the various *forms* of logic they have attained. However, it is a matter of historical record that children who failed on tasks were often simply described as non-logical, and that the problem of criteria of understanding has received relatively scant attention in Piagetian literature.

The representativeness of the tasks.

During my research work I gradually came to recognise that I presented the children with tasks that were entirely new to them and that were quite atypical compared with their normal everyday activities (Smedslund, 1966). The closest connection of these tasks with anything in children's lives was with mathematics and natural science as taught in the school systems of the technologically advanced world. It is commonly and only half jokingly said that even psychological researchers are afraid of going to Geneva because they fear they may fail on some of the tasks that the older children are given. The closed world of the school system with its competitive ideology of bookish knowledge and its emphasis on a certain conception of ' intelligence ' is in many ways controversial and may be accused of neglecting other important aspects of living. As an adjunct to this system, Piagetian psychology, therefore, has some potentially questionable *political* and *practical* implications. More specifically, it may be seen as supporting a one-sidedly intellectualistic view of life and as maintaining the prestige of abstract formalism. This is true even though Piaget's psychology also has contributed new impulses to education, especially by its emphasis on *constructivism.*

Working as a practising psychologist with children's problems in a real world, I have come to experience very strongly the irrelevance of much of my Piagetian background. Molar features of children's environment such as brutality, rejection, lack of contact, jealousy, lack of consistency, double communication, dependency, and in general all kinds of evil interactional circles dominate the picture and determine what action is to be taken. Performance and interest in school work and in theoretical tasks of the kinds favoured by Piaget, are linked in various ways with the child's family and cultural background but are also frequently strikingly unrelated to the degree of sophistication with which he behaves in his world outside the classroom. The touching and often quite complex philosophical reflections of children who fail at school and who would no doubt fail on the age-relevant Piagetian tasks, have made me very careful indeed when it comes to judging a child's capacity and potentialities. Actually, I think that the Piagetian tradition shares with the rest of academic psychology the unfortunate tendency to attribute success and failure on tasks to characteristics of the person as such instead of to characteristics of the person-in-his-total-life-situation (see Ichheiser, 1970, especially chapter 5).

In conclusion, the lack of representativeness of the tasks relative to the total *Lebenswelt* of the child is one of the reasons why Piagetian psychology is relatively useless for practical purposes. In addition, the bias towards abstract formalistic skills also has questionable political implications.

LASTING CONTRIBUTIONS TO PSYCHOLOGICAL PRACTICE

The preceding negative evaluations explain, at least partly, why I have moved away from the Geneva tradition. Yet, the picture would be very incomplete if I did not mention at least some of the teachings of Piaget which I have brought with me, and which I regard as valid and valuable. Before I do this, let me add that I do not pretend to be in any way comprehensive in my comments. Piaget obviously is one of the giants in the history of psychology and his contributions are of an order that cannot be dealt with adequately by one person in a brief article. Let me just summarise a few of the points that appear central from the point of view of psychological practice, in addition to the tremendous increments in factual knowledge of children, especially in Western societies in our time that Piaget has instigated.

(1) I cherish Piaget's view of the interplay between assimilation and accommodation and his constructivist position which is quite compatible with the *hermeneutic* spiral and with the conception of a *dialectical* psychology. It also implies an outright rejection of the still prevalent view that psychologists gather objective and theoretically neutral data.

(2) I believe that Piaget's support of the ancient '*méthode clinique*' made possible most of his contributions and was an important step away from the rigid objectivism of the traditional laboratory. It was also compatible with a recognition that the data of psychology are acquired through acts of *communication* with a shared meaning basis presupposed.

(3) Piaget's analysis of the notions of *centration* and *decentration*, especially as they apply to interpersonal relations, is undoubtedly a fundamental and lasting contribution. The fruitfulness of these concepts is especially clearly displayed in the recent analysis of Haavind and Hartmann (1976) of the strategies of mothers as teachers and of their children as learners.

(4) In my view, Piaget's emphasis on the mode of *implication*, and hence on *necessity*, *possibility*, *negation*, etc., is of crucial importance. It underlines the role of the *subject* in psychology and rejects behaviourism's view of man as *object*. Implication exists *for* someone, not as an objective fact.

(5) Finally, Piaget's accentuation of the role of *organisation* of the person's activity is profoundly important. This leads to the characterisation of activity in terms of strategies, rather than in terms of specific linkages between behaviour and elements in the environment. What characterises a child, an adult layman or a professional psychologist is not only the specific knowledge and specific skills they have, but also the way in which they have internalised, organised and reflected upon this knowledge and these skills. However, I do not agree with Piaget in describing organisation as approximations to formal logical structures, rather I think organisation of behaviour should be described concretely in terms of ordinary language and with no distinction between form and content. Examples of general strategies described in this way would be, for example, " in dangerous situations, think before you act," " in order to control a person's behaviour, manipulate the consequences of this behaviour," etc. Examples of therapeutic strategies described in ordinary language are given in Smedslund (1973).

CONCLUSION

I have tried to describe very briefly some main points in my current evaluation of Piaget's psychology. I probably owe more to Piaget than to any other single psychologist, yet I have become increasingly critical of his theoretical system. Working with children during my Piagetian period I felt constrained to be detached, one-sidedly cognitive and entirely focused upon certain very abstract aspects of performance. What mattered to me then were the formal logical structures rather than the concrete living children in their total life situations. I tended to ignore the artificiality of the tasks and I was oblivious of the political implications of my work as supporting a one-sidedly intellectualistic and school-centred ideology. I doubt that a practitioner can be a useful psychologist if he remains wholly within that tradition. Piaget has contributed a wealth of brilliant and penetrating insights, but they must be incorporated into a view of psychology which can be lived and practised rather than merely written and talked about in academic settings.

ACKNOWLEDGMENTS—I am indebted to Karsten Hundeide for many valuable comments.

REFERENCES

HAAVIND, H., and HARTMANN, E. (1976). Mothers as teachers and their children as learners. Unpublished manuscript, Institute of Psychology, University of Bergen, Norway.

ICHHEISER, G. (1970). *Appearances and Realities: Misunderstanding in Human Relations.* San Francisco: Jossey-Bass.

PIAGET, J. (1962). Introduction to LAURENDAU, M., and PINARD, A., *Causal Thinking in the Child. A Genetic and Experimental Approach.* New York: International University Press.

SMEDSLUND, J. (1966). Les origines sociales de la decentration. In BRESSON, F., and MONTMOLLIN, Mde. (Eds.), *Psychologie et Epistémolgie Génétiques.* Paris: Dunod.

SMEDSLUND, J. (1969). Psychological diagnostics. *Psychol. Bull.*, 71, 237-248.

SMEDSLUND, J. (1970). Circular relation between understanding and logic. *Scand. J. Psychol.*, 11, 217-219.

SMEDSLUND, J. (1973). The heuristics of interpersonal transaction: A strategy for group work. Work notes on small groups. (No. 8). Mimeographed Report, Institute of Psychology, University of Oslo.

WASON, P. C., and JOHNSON-LAIRD, P. N. (1972). *The Psychology of Reasoning: Structure and Content.* London: Batsford.

WASON, P. C. (1976). The theory of formal operations—a critique. Unpublished manuscript.

(*Manuscript received 15th October, 1976*)

Br. J. educ. Psychol., 47, 7-17, 1977

PIAGETIAN PSYCHOLOGY AND EDUCATION: TIME FOR REVISION

BY G. BROWN AND C. DESFORGES

(*Department of Educational Research, University of Lancaster*)

SUMMARY. The relationships between Piaget's theory and education are discussed. The accumulated data from studies based upon the theory are examined in terms of (i) shared assumptions which govern the interpretation of replication studies, (ii) the extent to which a stage can be meaningfully defined, and (iii) the nature of the developmental process proposed by the theory.

It is concluded that Piaget's search for generalities in cognitive development has been unsuccessful, and that an approach based upon a skill-integration model might be more profitable.

INTRODUCTION

DESPITE the popularity of Piagetian psychology in studies of education, the relevance of this theory to teaching has often been contested. Whilst Piaget's most convinced followers have been prepared to base whole curricula and teaching programmes upon it (Furth and Wachs, 1974), the most vociferous critics have taken the view that the theory has no implications whatsoever for instruction (Engelman, 1967, 1971). In general these contentious issues are rarely broached, and when criticisms are considered, there has been a tendency to trivialise them (as in McNally, 1974, for example). Yet the status of Piagetian theory is clearly a crucial issue for those educators who do derive pedagogic implications from it. The implications stem from three central notions:

(1) development occurs through an invariant hierarchy of stages, in which the successful negotiation of one stage is a prerequisite for the optimal development of the next;

(2) each stage has an underlying unity of operation, which applies to all intellectual skills exhibited within that stage; and

(3) the key process of development is identified as equilibration, which describes the process of reciprocal interactions between environmental 'inputs' (experience) and growing cognitive structures. This growth is partly due to maturation, but primarily arises out of the active interchanges between adaptation of the mental structures and variations in the intellectual content of environmental encounters.

The practices of assessing children's performances, sequencing curriculum materials, and structuring learning environments to facilitate progress are justified, not directly from Piagetian observations, but from the abstract notions of stages and developmental processes which Piaget takes as explanations of his observations. For example, the idea that the pre-operational child learns best by ' doing,' rests on Piaget's hypothetical stages and his pre-theoretical assumptions regarding the nature of intelligence. Since educational implications arise from the underlying principles of Piaget's epistemology, it is vitally important that educators understand this relationship between theory and practice, and that psychologists gather data which test the integrity of the system which is taken to authorise professional practice.

Basic educational psychology courses seldom have sufficient time for all those issues which the lecturers would wish to pursue. Consequently, simplified, second-hand sources are attractive. Unfortunately, such texts seldom address themselves to any criticisms of Piaget's theory other than those which have been termed ' undergraduate ' criticisms (Flavell, 1963). These focus upon the limitations of sampling and contingent limits of generalisations, the complexity of Piaget's instructions, and the arbitrariness of his scoring procedures.

Whilst these criticisms are not without substance, sufficient energy has been expended in elaborate replications with representative samples in a variety of cultures (see Elkind, 1961; Laurendeau and Pinard, 1970; Lovell, 1961; Smedslund, 1961) to make further reservations of this type unwarranted. It has been shown for example, that 5-year-olds generally fail to conserve number, whilst 6-year-olds succeed; and that most 11-year-olds are unable to accomplish tasks which require formal operations, such as the ' pendulum problem.'

There are, however, more serious criticisms which need consideration. They concern (i) the shared assumptions which have governed the interpretation of data in replication studies and in the original studies, (ii) the extent to which ' stage ' can be meaningfully defined and (iii) the nature of the developmental process. These are more weighty issues, and each must be scrutinised in turn.

SHARED ASSUMPTIONS IN REPLICATION STUDIES

We have already asserted that criticisms of Piaget's sampling procedures are now of minor consideration. Theorists who seek data only from the random sample, and eschew the insights of the clinician, cut themselves off from an important source of information. Nevertheless, irrespective of the source of an hypothesis, it is essential that at some point it is tested in order to eliminate alternative explanations. Piaget does not take this step. His data are presented as *consonant* with his theory. That is, it is illustrative, not confirmatory. Furthermore, those who seek to replicate his findings adopt the same analytic procedures. They confirm only that Piaget's data can be reproduced in other samples. They do not attempt to test the meaning which he ascribes to that data (see Wohlwill, 1968, for a discussion of this point).

Bryant's work (Bryant, 1974) examines this deficiency. Using Piaget's transitive deductive inference experiment, Bryant systematically investigated alternative explanations for the observations that after being presented with two relationships, A>B and B>C, children below the age of 8 years were often unable to make the inference A>C. The data are not contested; what is in question is Piaget's explanation that the cause of the phenomenon is the dominance of the immediate perceptual input, which is not amenable to reorganisation. Bryant proposed two alternative hypotheses: (*a*) that the children may be unable to recall the first instance presented, and (*b*) as the children had only seen A in a ' greater than ' position, they may be parrotting that in the new situation. In the latter case they would succeed, but for different reasons.

Systematic experiments, controlling these two variables, led to the conclusion that memory failure was certainly implicated in the results obtained from Piagetian protocols, although Bryant made clear that it was not likely to be the only variable involved, i.e., that memory and operativity are not independent. When memory failures were eliminated, 88 per cent of a sample of 5-year-olds were able to make the transitive inference. The importance of this finding lies not in the implications for changing the accepted age at which such operations occur, but in their demonstration of Piaget's failure to test alternative hypotheses.

Similarly, the work of Bower (1974) suggests that Piaget's interpretation of aspects of the sensori-motor stage is untenable if closer attention is paid to infant behaviour. According to Piaget, early sensori-motor behaviour shows no changes when objects which are in view are made to disappear, that is, there is no notion of object constancy. Later, at about 18 months, behaviour does show changes, because the object which has been seen can be represented in the child's mind by a ' signifier.' The advent of symbolism, in the form of this signifier, constitutes a major advance in development which was acclaimed by Elkind and Sameroff (1970) as ' the most significant of Piaget's discoveries.'

Bower's evidence suggests that observers who had expected to find no change in behaviour saw no change in behaviour. He expressed no surprise that very young infants did not reach out for objects or crawl after them, but closer examination suggested that there were indicators of change. In an ingenious set of experiments using objects which were alternatively visible and occluded, Bower's recordings of head and eye movements and changes in heart rate indicated that infants anticipated the re-emergence of occluded objects and registered surprise at their non-reappearance. As in Bryant's work, systematic experimentation enabled alternative explanations to be eliminated. Bower concluded that " these data . . . would seem to prove conclusively that quite young infants . . . know that objects still exist after they have been occluded . . . " (p. 189). So one of Piaget's ' significant discoveries,' which constitutes an important item in explanations of cognitive development in the second year of life, is now thought to be present at a very much earlier time.

THE DEFINITION OF A STAGE

It is generally agreed that the idea of ' stage ' involves a coherent integration of operations into a theme, or a series of themes, which underlie certain categories of behaviour at a particular point in time. In Piagetian terms this orchestration of themes is a *structure d'ensemble.* To the best of our knowledge no one has been sufficiently naive to suggest that stages change by means of some mystical transformation which alters or replaces these themes overnight. Consequently, we expect that at times of transition an operation characteristic of a particular stage might not be manifest on all apparently appropriate occasions. These lapses of performance are described as examples of *horizontal décalage.* Nevertheless, if the stage concept is to have utility, there must be a substantial degree of homogeneity within the behaviours of most children for the greater part of each stage. If the extent of *décalage* exceeds that of *structure d'ensemble,* in other words if heterogeneity is characteristic of most children's performances most of the time, we are unable to make predictions about behaviour or make provision for educational tasks except in those specific areas which we have actually assessed. In these circumstances it behoves educational psychologists to seek other explanations of their observations.

Considerable numbers of studies have been accumulating, casting doubt upon the integrity of stages. Some refer to the surprisingly low correlations between behaviours at a given time, others to the absence of expected operations, and yet others to the presence of operations commonly taken as representative of much later stages.

There has been a succession of studies reporting low correlations between various measures of a particular operation (see Pascual-Leone, 1970, for example). More recently, Hamel (1974) found correlations of from ·39 and ·72 between multiple classification scores and multiple seriation scores. Whilst these correlations are significant, Hamel raises the question of how high one should expect them to be if they are indicative of a *structure d'ensemble*. In studies relating, say, different sub-tests of some test of abilities, such figures would be highly respectable, but as indicators of the relationships within a unified stage of cognitive behaviour it is surprising that none accounts for more than about 50 per cent of the variance.

In a further study Hamel gave pre-operational tasks, comprising 240 items, to children of 5½ and 9 years. The items were concerned with groupings of classification, seriation, multiple classification and multiple seriation. Analyses of variance showed that there was very considerable significant variation both within and between groupings. Furthermore, there was considerable evidence of task-specificity in the children's responses, with the greatest number of correct answers given for tasks of classification, followed by seriation, multiple classification, and multiple seriation respectively. The specificity was heightened by significantly more correct solutions being given if the task involved judgment on a colour dimension, and least correct solutions if number was involved. This study very clearly suggested that there was little evidence of a *structure d'ensemble* for any of the children in Hamel's sample.

Similar evidence of heterogeneity is found in the work of Piaget's own associates. In an experimental manipulation of training procedures for the conservation of continuous quantity (Inhelder *et al.*, 1974) 34 children between the ages of 5 and 7 years were given tasks involving conservation of liquids and modelling clay. Fifteen children were categorised as ' non-conservers ' but 19, more than half the group, were euphemistically labelled ' intermediate.' Of those, six failed the first task but succeeded on the second, nine gave mixed correct and incorrect answers on both tasks, and four were successful on the first but not the second task. Similar margins of ' intermediacy ' are to be seen in most of the data they present.

Other studies show similar results. Data reported by Neimark (1975) were concerned with her extensive longitudinal study of formal operational thinking. Piagetian tests of permutation and combination were administered on several occasions as children developed from grades 3–6 to 7–10. Test-retest correlations and between-measures correlations were all significant, but they were of the order of 0·4. This raises the question already posed with reference to Hamel's study. Is a correlation which accounts for so little variance really sufficient to substantiate the existence of a *structure d'ensemble*? Of the 58 children in Neimark's sample who scored at ' level 3 ' on the combinations test, 24 scored at ' level 0 ' on the permutations tests, eight at ' level 1,' 19 at ' level 2,' and seven at ' level 3.' Clearly, there is little evidence of concurrent development of these formal operational skills, as Neimark noted.

Performance on Piagetian tests by subjects who could be taken to be fully within the formal operational stage have been reported by Schwebel (1975). His subjects were 30 men and 30 women, with a mean chronological age of 19 years, who were, on average, one standard deviation above the mean on tests of mathematical and mental abilities. He presented them with rods, balance scales and truck problems (Piaget and Inhelder, 1958). The mean intercorrelation between performances on these tests was 0·33, the highest being 0·37. Schwebel concluded that " the three tasks appear to be relatively independent of each other " (p. 137).

This diversity of performance on tests which supposedly measure aspects of the *structure d'ensemble* of formal operations, is evident in the extensive review, which Blasi and Hoeffel (1974) produced, of a large number of studies similar to those cited above. The percentage of ' average ' children between the ages of 11 years and 14 years judged to be formal operational by various tests varied from 100 per cent to zero. Even when the same test was applied to different samples the percentage varied from 0 to 35 per cent.

In 1972, Wason and Johnson-Laird reported evidence of heterogeneity being more impressive than that of homogeneity on tests involving formal, or logical, problems. In some of their studies, subjects were presented with problems which had either a concrete or an abstract content. Even those subjects who succeeded on the concrete problem generally failed to transfer their solution to the same problem couched in abstract terms. This led the authors to conclude that formal operations were " not cognitive skills which can be applied to any problem, they are really practical rules rather than formal operations " (p. 190).

That studies critical of formal operations are received without much surprise is probably due to the strong reservations which many psychologists hold concerning the adequacy of Piaget's description of this stage (Lunzer, 1974, for example). It is tempting to speculate that, when faced with a description of the developmental stage in which psychologists find themselves categorised, the shortcomings of the model are all too apparent. Unfortunately, time and the parsimony of memory have made us receive, all too readily, an oversimplified model of cognitive processes in childhood.

The exact proportion of heterogeneity which theorists will tolerate without abandoning the stage concept, is an interesting point. The extreme provocation is presumably presented by evidence of operations considerably in advance of those expected, and which are undetected by standard Piagetian procedures, for we are then obliged to ask whether many other instances of sophisticated behaviours might exist. Both Bryant and Bower gave evidence of ' precocious ' behaviour. Similar results were found by Povey and Hill (1975). Fifty-six children between the ages of 2·4 and 4·10 years were given concept acquisition tasks involving picture cards, and half of them were asked some Piagetian questions. The authors found that these questions had little discriminatory power, but their own material indicated that the children had formed specific concepts. This was particularly true of the older children. Povey and Hill stressed that the key point in their disagreement with Piagetian theory was not that the age-stage relationship was incorrect, although that seemed to be the case, but that the criterion tests which have been accepted as indicators of class inclusion are inadequate. At their least damaging, these findings suggest that class inclusion does *not* belong in the stage of concrete operation, but in preoperational thinking. But we do not know how many other Piagetian tests are telling us that certain operations are absent, when better tests would tell us that this is not so. So the argument is not a skirmish over fine detail, but a forceful attack on the entire structure of stages, backed up by other evidence of ' precocity ' (Harris, 1975; Gelman, 1972; Gelman and Tucker, 1975).

Harris (1975) explored the abilities of children of 5 to 7 years to infer the attributes of nonsense concepts from knowledge of the class membership of the concept. He concluded that children of this age appeared to possess the necessary operations for inference, and suggested that the definition of the ' formal operational stage ' may need reformulating in terms of the " spontaneous deployment of rules of inference rather than their availability as such " (p. 151).

' Precocity ' was also reported by Gelman (1972). Children between the ages of 2½ and 4 years were found able to conserve pattern recognisable numbers, understand quantity changes caused by addition and subtraction, and understand the ' one unit ' difference between the numbers 2 and 3. In a later study (Gelman and Tucker, 1975), children of 3, 4 and 5 years were able to discriminate colour changes from number relevant changes in arrays of three items. All these observations are contrary to expectations based on Piaget's definitions of the operations present in specific stages.

Piaget suggests that a child enters a new stage by evolving certain task-specific operations which, over a period of time, converge to form a more generalisable *structure d'ensemble* of the new stage. He admits that certain tasks appear ' resistant ' to the new operations, so that the process of change and the peculiarities of some tasks are causes of observed *décalage*. The point has been made that it is necessary to decide how widespread this *décalage* can be without it seriously eroding the utility of the stage model. We have tried to show that it is very widespread indeed; but it may be that all the studies to which we have alluded have concentrated on the specific transitional phases, when heterogeneity is to be expected. In this case we would expect to find expanses of the childhood time-scale which these studies have ignored, and in which it would be reasonable to expect more consistent behaviour. Bower's study investigated infants between the ages of 2 and 6 months; Povey and Hill, and Gelman used children between 2 and 5 years; Inhelder *et al.*, Hamel, Bryant, and Harris all employed children of between 5 and 7 years; Neimark's subjects were aged 9 to 16 years; Blasi and Hoeffel's were between 11 years and adulthood, and Wason and Johnson-Laird's subjects were all adult students. The evidence seems to speak for itself.

In rejecting Piaget's description of stages, there is no implied rejection of the notion of some sort of coherence or system in cognitive growth. Behaviourist attempts to explain transfer of training, in terms of the observable characteristics of the tasks, foundered, not because there was no evidence of systematic transfer, but because it became apparent that the systematisation seemed to be subjective in nature. Intellectual growth is systematic, but the stage descriptions Piaget offers as evidence of the system seem less than adequate. The next step in the argument is to examine his description of the dynamics of his system.

THE NATURE OF DEVELOPMENTAL PROCESSES

The Genevan school rejects the view that either empiricism alone or nativism alone can account for the dynamics of the acquisition of knowledge. The limitations of these philosophies are stated throughout Piaget's work and have been most recently and convincingly expressed by Inhelder *et al.* (1974) and Furth (1974). Thus, whilst acknowledging the importance of underlying maturational processes and the necessity for social and physical experience, Piaget and his colleagues carry " the epistemological conviction that progress in human knowledge results from dynamic processes which imply self-regulating mechanisms . . . " (Inhelder *et al.*, 1974, p. 242). This conviction manifests itself in their proposal of an epigenetic system in which the source of progress is the persistent clash between the modes of intellectual processing available at a certain stage of development and certain experiences. " The source of progress is to be sought in the disequilibrium which incites the subject to go beyond his present state in search of new solutions " (Inhelder *et al.*, 1974, p. 264). Through these processes of disequilibrium and subsequent equilibration, new intellectual structures are developed which incorporate the preceding structures, and consequently development is through an invariant sequence of hierarchically ordered stages. Hence, whilst maturation and various social and physical experiences are considered to be essential to development, the central dynamic process is that of equilibration.

A number of educational implications is taken to arise out of this view of development. It is taken to support the view of learning by self-paced discovery methods (Yardley, 1970), to justify activity methods in learning (Almy *et al.*, 1966) and to warrant a readiness stance with respect to the use of formal teaching methods (Isaacs, 1965; Badcock *et al.*, 1972; Ross *et al.*, 1975). With this in mind it seems important to identify the status of the concepts of 'equilibration' and the 'hierarchically organised epigenetic system.' In particular we would want to know whether these concepts are based on data or whether they are merely based on 'epistemological conviction.'

In earlier criticisms of Piaget's work these functional concepts were viewed with some concern. They were seen by Flavell (1962) as very broad generalisations containing "only the seeds of process rules" (p. 85) and by Bruner (1959) as contributing nothing to the theory. Sullivan (1967) contended that this was the weakest part of the theory both in terms of specificity of description and scientific testability. Most recent work in this context emanating from Geneva confirms both the centrality and necessity of the notion of equilibration together with a continuing inability to specify its detailed functioning or to pursue its empirical exploration. As a consequence, examination of the data pertinent to this issue has to be somewhat tangential.

With respect to the role of maturation and the universality of cognitive structure, it is claimed that cross-cultural studies have confirmed the consistent order of acquisition of various Piagetian concepts (Inhelder *et al.*, 1974, p. 246). However, in a review of cross-cultural Piagetian research, Ashton (1975) is not nearly so confident. It seems that whilst sequential development has been confirmed in Western studies, contradictory evidence is available from Thai, Arab, Indian, Somali and Australian Aborigine children. Ashton is cautious about the interpretation of these studies since their comparison involves problems of methodology, materials and the shared assumptions about the transferability of questioning techniques across cultures. Even studies which find similarities of sequence but with marked developmental lag, i.e., operations appearing much later in one culture than in another, offer only 'neutral' evidence. That is, they fail to discriminate between the effects of maturation and experience. Some studies show that particular experiences and particular task requirements interact to make any generalisations hazardous. It has been observed that children of Mexican potters perform better on conservation of substance tests than children of 'non-potters' (Price-Williams *et al.*, 1969). Greenfield (1966) demonstrated the positive effects of schooling on the acquisition of conservation, whilst Maccoby and Modiano (1966) suggested that unschooled children were superior. Kohlberg (1968) found that amongst a Malaysian aboriginal group conservation of substance was attained at 7–8 years but partially lost between 11–15 years, due, Kohlberg hypothesised, to conflict with adult magical beliefs. These data do not refute the need for an epigenetic model but emphasise the difficulties of identifying such a model.

We have already shown the role of experience to be crucial in determining the outcomes of particular tasks. This has been emphasised in numerous training studies. The central questions in these studies have been (*a*) what are the limits of development that such experience can facilitate and (*b*) are their 'optimal' experiences? Once again the data are extremely difficult to interpret. Engelman has claimed he can train 5-year-olds to do formal operational problems (Engelman, 1967, 1971), but his claim has been closely criticised by Kamii and Derman (1971). In the main the dispute is less about children's post-training performance than about the failure of such training studies to demonstrate some generalisable persistent gain which is resistant to extinction. The failure of a great many training studies to monitor gains after post-training makes them rather useless in any attempt to evaluate quite what the training did in itself. The Piagetian school, however, recognises the value of properly conducted training studies and, indeed, sees them as necessary if we are to identify the dynamics of development. Inhelder's recent review of Genevan attempts in this direction (Inhelder *et al.*, 1974) show them to be a long way from realising their aims. Furthermore, these studies show that ' experience ' does not influence all children equally. Children in transitional states prior to training subsequently show superior progress to that of children from the same age not in the transitional states, i.e., experiences can indeed be ' optimal.' However, there is nothing in empiricism that prevents it accounting for the differential effects of experience.

It seems that cross-cultural and training studies presently available do not illuminate our understanding of the dynamics of development. This is only in part due to methodological limitations and problems. More centrally it is due to the fact that such studies have rarely concerned themselves with identifying and testing fundamental epigenetic processes of the theory.

Brainerd (1973) cites a number of exceptions to this general failure. A more recent example is reported by Bower (1974). To compare a simple learning model with Piaget's 'reconstruction through conflict' model of development, Bower subjected two groups of 80-day-old infants to training schedules to examine the effects of this training on their performance on the Mundy-Castle task. In this task infants tracked an object as it appeared consecutively in each of four portholes in the order, bottom left, top left, top right and bottom right. At 12 weeks the infants were able to anticipate the appearance of the object but only by simple eye movements such as place to place tracking. A little later the infants' tracking began to demonstrate interpolation of the trajectory between portholes. Bower's two groups of infants initially demonstrated the simple tracking behaviour. One group was given circular tracking training (identical in operation to the more sophisticated performance of the Mundy-Castle task) whilst the second group was given oscillation tracking training. Given the hypothetised 'rules of tracking' of children this age, 'oscillation' training is rich in conflict. In simple behavioural terms, however, it is not advanced on the simple tracking behaviour initially deployed on the Mundy-Castle test. A simple behavioural model of development would predict superior transfer of training for the first group to the Mundy-Castle task. The Piagetian model would predict that the second group would show significantly better facility on this task because of development arising out of the conflict-provoking training schedule. In fact, the Piagetian prediction was confirmed. The second group showed four times as many trajectory interpolations as the first group on post-training tests.

From this evidence it does appear that Piaget's emphasis on conflict as an important feature of the equilibration process may be justified. It is equally clear, however, that there is very little information about the all-important processes which lie at the heart of his theory. In consequence, educational implications arising from the epigenetic model are based almost exclusively upon ' epistemological conviction.'

A POSSIBLE ALTERNATIVE MODEL

The pedagogical questions raised earlier referred to the problems involved in sequencing and structuring curriculum materials, and in presenting learning experiences appropriate to a notional developmental level of the learner. The Behaviourists see these problems as totally located in the logic of the material itself, and the learner's position in relation to the sequence of the material. Piaget attempts to describe structure in terms of the learner's mental operations, which are postulated as underlying intellectual performance. We have already alluded to the necessity of a structuralist model, but our survey of Piagetian studies leads us to conclude that the structure this theory proposes has not been substantiated. In countering the specificity of behaviourist models, it has strayed too far into generality. The search for a structure irrespective of content has led to generalities which cannot be sustained. More of the same kind of research is unlikely to ' fill the gaps,' for there is no paucity of information, but an embarrassment of dissonant information. The vague conclusions of Piagetian research may well be a consequence of the search for content-free mental structures, a strategy which is arguably misguided. It would seem more profitable to locate cognitive structures within specific content domains.

Several recent studies suggest a move in this direction: Bruner (1973), Klahr and Wallace (1973) and Schaeffer *et al.* (1974). Each of these studies adopts a ' skill integration ' model of development. Within this model skills are taken to be procedures, or sets of procedures, deployed *in particular settings*. Integration of these skills involves increasing generalisation to new situations, increasing automation (i.e., less demand on the processing capacity of focal attention), and the development of schemes to unite various skill elements.

Schaeffer *et al.* (1974) describe a skill integration model for the development of number conservation. The counting behaviour of children between the ages of 2·0 and 5·11 years was investigated. Children were asked to count drum taps; to put ' x ' sweets in a cup ; to drum-tap up to ' x ' ; to pattern-recognise small numbers; to count objects and report the number after they were covered up; to choose between ' x ' and ' x+1 ' sweets; to count objects in ordered arrays, disordered arrays, with sub-groups and without sub-groups, with pointing and without pointing. By examining the order in which these tasks were successfully completed, it was possible to organise them into a skill hierarchy. The model accommodates Gelman's data (1972) on the abilities of very young children to conserve pattern-recognisable numbers, whilst at the same time taking into account the serious contextual and set-size limitations of these abilities as demonstrated by traditional Piagetian studies. In Schaeffer's terms conservation involves the integration of three major subskills: first, the capacity to judge relative numerosity in small sets; second, the capacity to deploy one-to-one correspondence; and finally, the acquisition of the cardinality rule (i.e., that the last number named when counting an array is the number of things in that array). Each of these subskills can, in turn, be specified in terms of more limited component subskills.

The hierarchical organisation of subskills carries the structuralist element of the model. The contextual limitations are recognised in two ways. On the one hand, no claim is made regarding the capacity of this particular skill hierarchy to describe the intellectual behaviour of 6-year-olds with other material. On the other hand, the origins of the number conservation hierarchy are not to be found within the general behaviour of the child, but emerging from quite specific experiences. The model " assumes that number development is determined more by the application of number skills to object arrays than by thoughts about number skills in the absence of arrays, or by spontaneous cognitive reorganisations, i.e., equilibration " (Schaeffer *et al.*, 1974, p. 378).

If this model is a possible replacement for the Piagetian model, we need to assess it on the same criterion: namely, what does it offer to the educator? To answer this we shall need more information, particularly regarding the generalisability of the skills. Are we to expect a hierarchy of number skills to operate on the quantification of area or angles? Or will each require a new set of skills ? In a sense we are revisiting the Behaviourist problem of what constitutes stimulus similarity? The important difference is that skill integration theory operates on two dimensions, that of stimulus identity (content) and cognitive processes (structures of skills).

It would be mischievous to ignore the problems associated with this model. The educator's concern is with how children perform specific tasks, and his aspiration is to predict how tasks may be presented for optimal learning. Yet the model allows only that, given an array of tasks, the presence of hierarchical rules can be determined. The implications for curriculum design are that, initially, we must abandon the search for general structures and set about producing taxonomies of behaviours for specific areas of the curriculum. In the long term, when sufficient taxonomies have been established, we may look again for general structures, but to start, as Piaget does, with a search for such generality has proved to be inappropriate.

REFERENCES

Almy, M., Chittenden, E., and Miller, P. (1966). *Young Children's Thinking.* New York: Teachers' College Press.

Ashton, P. T. (1975). Cross-cultural Piagetian research: an experimental perspective. *Harvard Educ. Rev.*, 45, 475-506.

Badcock, E. H., Daniels, D. B., Islip, J., Razzell, A. G., and Ross, A. M. (1972). *Education in the Middle Years.* Schools Council Working Paper No. 42. London: Evans/Methuen.

Blasi, A., and Hoeffel, E. C. (1974). Adolescence and formal operations. *Human Dev.*, 17, 344-363.

Bower, T. G. R. (1974). *Development in Infancy.* San Francisco: Freeman.

Brainerd, C. J. (1973). Neo-Piagetian training experiments revisited: is there any support for the cognitive-developmental stage hypothesis? *Cognition*, 2, 349-370.

Bruner, J. S. (1959). Inhelder and Piaget's 'The growth of logical thinking'—I A psychologist's viewpoint. *Br. J. Psychol.*, 50, 363-370.

Bruner, J. S. (1973). Organisation of early skilled action. *Child Dev.*, 44, 1-11.

Bryant, P. (1974). *Perception and Understanding in Young Children.* London: Methuen.

Elkind, D. (1961). The development of quantitative thinking: a systematic replication of Piaget's studies. *J. genet. Psychol.*, 98, 37-46.

Elkind, D., and Sameroff, A. (1970). Developmental psychology. *Ann. Rev. Psychol.*, 21, 191-238.

Engleman, S. (1967). Teaching formal operations to pre-school advantaged and disadvantaged children. *Ontario J. educ. Res*, 9, 3.

Engelman, S. (1971). Does the Piagetian approach imply instruction? In Green, D. R., Ford, M. P., and Flamer, G. B. (Eds.), *Measurement and Piaget.* New York: McGraw-Hill.

Flavell, J. H. (1962). Discussion of paper by Kessen. In Kessen, W., and Kuhlman, C. (Eds.), Thought in the young child. *Monog. Soc. Res. Child Dev.*, 27, 65-82. Whole of No. 83.

Flavell, J. H. (1963). *The Developmental Psychology of Jean Piaget.* London: Van Nostrand.

Furth, H. G. (1974). Ontogeny: development and learning. In Reese, H. W. (Ed.), *Advances in Child Development and Behaviour: Vol.* 9. London: Academic Press.

Furth, H. G., and Wachs, H. (1974). *Thinking Goes to School.* New York: Oxford University Press.

Gelman, R. (1972). The nature and development of early number concepts. In Reese, H. W. (Ed.), *Advances in Child Development and Behaviour: Vol.* 7. London: Academic Press.

Gelman, R., and Tucker, M. F. (1975). Further investigations of the young child's conception of number. *Child Dev.*, 46, 167-175.

Greenfield, P. (1966). On culture and conservation. In Bruner, J., Olver, R. R., and Greenfield, P. (Eds.), *Studies in Cognitive Growth.* New York: Wiley.

Hamel, B. R. (1974). *Children from* 5-7. Rotterdam: University Press.

Harris, P. (1975). Inferences and semantic development. *J. Child Lang.*, 2, 143-152.

INHELDER, B., SINCLAIR, H., and BOVET, M. (1974). *Learning and the Development of Cognition*. London: Routledge and Kegan Paul.

ISAACS, N. (1965). Memorandum for the Plowden Committee. *Froebel J.*, 2, 12-33.

KAMII, C., and DERMAN, L. (1971). Discussion of Engleman's paper. In GREEN, D. R., FORD, M.P., and FLAMER, G. B. (Eds.), *Measurement and Piaget*. New York: McGraw-Hill.

KLAHR, D., and WALLACE, J. G. (1973). The role of qualification operators in the development of conservation of quantity. *Cogn. Psychol.*, 4, 301-327.

KOHLBERG, L. (1968). Early education: a cognitive-developmental view. *Child Dev.*, 39, 1013-1062.

LAURENDEAU, M., and PINARD, A. (1970). *The Development of the Concept of Space in the Child*. New York: International University Press.

LOVELL, K. (1961). A follow-up study of Inhelder and Piaget's 'The Growth of Logical Thinking.' *Br. J. Psychol.*, 52, 143-154.

LUNZER, E. A. (1974). Formal reasoning: a re-appraisal. Unpublished paper. University of Nottingham.

MACCOBY, M., and MODIANO, N. (1966). On culture and equivalence—I. In BRUNER, J., OLVER, R. R., and GREENFIELD, P. (Eds.), *Studies in Cognitive Growth*. New York: Wiley.

MCNALLY, D. W. (1974). *Piaget, Education and Teaching*. Lewes: New Educational Press.

NEIMARK, E. D. (1975). Longitudinal development of formal operations thought. *Genet. Psychol. Monogr.*, 91, 171-225.

PASCUAL-LEONE, J. (1970). A mathematical model for the transition rule in Piaget's developmental stages. *Acta psychol.*, 32, 301-345.

PIAGET, J., and INHELDER, B. (1958). *The Growth of Logical Thinking from Childhood to Adolescence*. New York: Basic Books.

POVEY, R. M., and HILL, E. (1975). Can pre-school children form concepts? *Educ. Res.*, 17, 180-192.

PRICE-WILLIAMS, D. R., GORDON, W., and RAMIREZ, R. M. (1969). Skill and conservation: a study of pottery-making children. *Dev. Psychol.*, 1, 769.

ROSS, A. M., RAZZELL, A. G., and BADCOCK, E. H. (1975). *The Curriculum in the Middle Years*. Schools Council Working Paper No. 55. London: Evans/Methuen.

SCHAEFFER, B., EGGLESTON, V., and SCOTT, J. L. (1974). Number development in young children. *Cogn. Psychol.*, 6, 357-379.

SCHWEBEL, M. (1975). Formal operations in first-year college students. *J. Psychol.*, 91, 133-141.

SMEDSLUND, J. (1961). The acquisition of conservation of substance and weight in children: I, II, III, IV, V, VI. *Scand. J. Psychol.*, 2, 11-20, 71-84, 85-87, 153-155, 156-160, 203-210.

SULLIVAN, E. (1967). *Piaget and the School Curriculum*. Bulletin of the Ontario Institute for Studies in Education.

WASON, P. C., and JOHNSON-LAIRD, P. M. (1972). *Psychology of Reasoning: Structure and Content*. London: Batsford.

WOHLWILL, J. F. (1968). Piaget's system as a source of empirical research. In SIGEL, I. E., and HOOPER, F. H. (Eds.), *Logical Thinking in Children*. New York: Holt, Rinehart and Winston.

YARDLEY, A. (1970). *Discovering the Physical World*. London: Evans.

(*Manuscript received 9th June, 1976*)

Br. J. educ. Psychol., 47, 18-24, 1977

ACTIVITY IN THE LEARNING OF PIAGETIAN OPERATIONAL THINKING

By W. S. ANTHONY
(*School of Education, University of Leicester*)

SUMMARY. Piaget has stated that the child's active manipulation of objects is a necessary part of teaching for intellectual development. But this appears to be contrary to empirical evidence. From a review of literature, it is concluded that Piaget's insistence on physical activity, and his dismissal of observation, are excessive.

THE purpose of this paper is to argue that a hypothesis about teaching, derived from Piaget's theory, is false. The main practical implication drawn from Piaget's theory by many writers, is the insistence on physical activity for intellectual development. The point is made by Flavell (1963, p. 367), by Hooper (in Sigel and Hooper, 1968, p. 429), by Ginsburg and Opper (1969, p. 221), by Beard (1969, pp. 72-75, 92-96), and by Piaget (1974) himself. In Piaget's theory, the term 'concrete operation' does not mean that the child must perform a physical operation every time he 'conserves' or classifies or makes a transitive inference, but nevertheless, the theory is that these types of 'concrete operational' thinking about objects have developed from prior physical activities upon objects. The child's physical manipulation of concrete material is supposed to be a necessary part of any teaching designed to lead the child toward concrete-operational thought.

Piaget has found that many people have misunderstood him on this point, and, therefore, it is worthwhile to quote him at some length:

> " . . . it appears that many educators, believing themselves to be applying my psychological principles, limit themselves to showing the objects without having the children manipulate them, or, still worse, simply present audio-visual representations of objects (pictures, films, and so on) in the erroneous belief that the mere fact of perceiving the objects and their transformations will be equivalent to direct action of the learner in the experience. The latter is a grave error since action is only instructive when it involves the concrete and spontaneous participation of the child himself with all the tentative gropings and apparent waste of time that such involvement implies. It is absolutely necessary that learners have at their disposal concrete material experiences (and not merely pictures), and that they form their own hypotheses and verify them (or not verify them) themselves through their own active manipulations. The observed activities of others, including those of the teacher, are not formative of new organisations in the child " (Piaget, 1974, pp. ix-x).

Thus Piaget agrees with the commentators cited above and emphasises that the child's active manipulation of objects is a necessary (though not sufficient) part of teaching for intellectual development. Nevertheless, Piaget also implies that this recommendation is a hypothesis which " must be verified, through actual classroom practices and educational results, rather than merely based on simple deductions " (Piaget, 1974, p. ix).

But, of course, hypotheses may, like the child's hypotheses above, be tested and *not* verified. The Piagetian hypothesis (that the child's manipulation of objects is a necessary part of teaching for intellectual development and that observation of others is not instructive in this respect) is logically more open to falsification than verification, and indeed, there is empirical evidence that the hypothesis is false. It is *not* necessary for the teacher to get the child to manipulate objects, and observation of others *is* valuable for intellectual development.

Thus, Anthony (1972, p. 76) has already pointed out that many experiments have had some apparent success in teaching concrete operational thinking *without* involving physical manipulation by the child in the training, for example, the experiments by Kohnstamm (1963), Gruen (1965), Beilin (1965), Beilin *et al.* (1966), all of which were reported in Sigel and Hooper (1968). More recent results of this kind have been reprinted by Hamel and Riksen (1973) and Brainerd (1974). In all these experiments, the descriptions of the training procedures imply that the teacher, not the child, manipulated the material. From Piaget's previously-quoted point of view, this type of training should have been useless for the development of operational thinking. Now, it could be objected that the child might, unrecorded, play with the material, or with similar material outside the training sessions. But even if such activity does take place, the fact remains that the training procedures were not *designed* to allow for it. Obviously, a training procedure is not going to specify that a child be bound in chains or drugged into paralysis. The point is that the training procedures have been found to be useful even though the child is *not* encouraged or instructed to act on the material (" with all the . . . apparent waste of time that such involvement implies," Piaget quoted above).

But the apparent success of these training procedures must now be examined more closely. The methods used in the evaluation of Piagetian training are disputable. Vernon (1976, p. 32) remarks that the results of the research " seem too complex to justify easy generalisations "; Kuhn argues that the results are " ambiguous and inconclusive " (1974, p. 590) ; Brown and Desforges (1977) also stress the failings of such research. Nevertheless, in the present paper, it is argued that researchers using *different methods* of evaluation, have obtained *similar results*, so that after all a positive and generally acceptable conclusion emerges.

It is necessary to review the problem of evaluation briefly, so as to explain the differences between the various experiments. The problem of the evaluation of training in concrete operational thinking is largely one about what kinds of *transfer* of training (or ' generalisation ') ought to be required, if the training is to be regarded as successful. " Thus, for any investigator who pronounces his study a success, another investigator may point to any absences of generalisation and claim that subjects didn't really attain the concept. Until the generalisation question is satisfactorily answered, then, pronouncing any given training study a success remains equivocal " (Kuhn, 1974, p. 593). Piaget and Inhelder have written: " . . . at the level of concrete operations logical forms are not yet independent of their content. They are a structurisation of the particular content and there is no necessary generalisation " (1969, p. 170). This implies that one should not necessarily expect transfer from class inclusion with beads to class inclusion with flowers. However, it is usual to test for this kind of transfer. Another form of transfer arises from the idea of the *structure opératoire d'ensemble* which is theoretically characteristic of the concrete-operational stage as a whole (but see Brown and Desforges, 1977). According to this idea, one expects transfer of learning between conservation, class-inclusion, and transitivity. However, this kind of transfer has rarely been tested. Further aspects of transfer are in the often-required explanations of conservation judgments, and in the sometimes-required resistance to counter-suggestion. These should be kinds of transfer of training inasmuch as, if the child merely emits explanations and resistances which he has already heard or practised, these performances will have little or no intellectual significance. A further, and often-used, criterion of successful training in concrete operational thinking is durability of a performance over time.

The experiments cited above will now be described very briefly to display the nature of the success achieved by the training. Kohnstamm (1963) developed procedures for training children in class-inclusion, and found transfer of learning to the same type of task but with different material. The children were still successful two to three weeks later. Inhelder and Sinclair (1969) repeated Kohnstamm's procedures and obtained similar results, but they added more post-tests demanding transfer to somewhat different class-inclusion tasks; the children did not succeed well on these different tasks and Inhelder and Sinclair concluded that ' true ' operational structures were not acquired.

Gruen (1965) trained children in number conservation and found transfer of learning to the same task with different objects. In the post-test, the children were apparently able to give good explanations (of their judgments) which they had not been told in training, though this is not quite explicit in the report. There was some transfer to length and substance conservation.

Beilin (1965) trained children in conservation of number and of length. The post-tests were typically two to three weeks after the end of training. The training transferred to similar materials, and the children learned to give correct statements of the principle of conservation. But their judgments may be attributed to stimulus generalisation, since the training and test arrangements were perceptually similar, the correct responses were those practised in training, and the enunciated verbal principle could have been simple repetition of what the children had been told. The training did not transfer to different conservation tasks. Beilin *et al.* (1966) trained children in the representation of water-levels (horizontality). The success of the training may be attributed to a limited stimulus generalisation; the training did not transfer to a slightly different shape of flask.

Hamel and Riksen (1973) trained children in the conservation of continuous and discontinuous quantity. The training transferred to a different form of test and to different dimensions (two-dimensional space, number, substance, weight). The effects were still evident one week later. The training required the child to make correct judgments of equality, but the tests required correct judgments of both equality and of inequality, which implies that the children did not simply learn a taught response. Explanations of equality judgments were also required, but it seems that these could have been simple repetitions of taught explanations.

Brainerd (1974) trained three different groups of children in transitivity, conservation, and class-inclusion respectively. The post-tests were one week after training. Each type of training was (partially) successful with respect to its type of task; there was also transfer between dimensions, from length to weight, including transfer to different materials. Some apparent transfer from transitivity to conservation was thought possibly due to errors of pre-testing. Other transfer between transitivity, conservation, and class-inclusion, was tested but not found. The control groups in this experiment received treatments similar to the trained groups except that the trained groups were given verbal correction of their answers whereas the control groups were not; the results so far mentioned are, therefore, attributed to the effect of verbal correction. However, it seems that the control treatments were (somewhat) effective training procedures in themselves, so that the subtraction of their results from the results of the training treatments is likely to be an unfair handicap on the latter.

Each of the experiments resulted in some transfer of training, but this does not end the argument. It was noted that Inhelder and Sinclair (1969) maintained that Kohnstamm had not demonstrated *sufficient* transfer. Inhelder and Sinclair, and later, Inhelder, Sinclair and Bovet (1974), set out principles for the evaluation of Piagetian training. In view of Inhelder's status as a long-standing colleague of Piaget, these principles may be taken as closely representing the Piagetian tradition. None of the experiments previously cited fulfils all the conditions recommended by Inhelder and co-authors. The evidence cited above, in support of the conclusion that the child's physical manipulation of objects is not a necessary part of training for concrete operational thinking, would perhaps be rejected by Inhelder and other Piagetians. Nevertheless, there is no deadlock here; for although Inhelder, Sinclair and Bovet might reject the *evidence* so far cited, they agree with the *conclusion*. Although these authors follow Piagetian theory in general, the statement below will be found to be *opposed* to Piaget's assertion of the necessity of the child's actual manipulation of objects:

> " Cognitive development results essentially from an *interaction* between the subject and his environment. In terms of successful training procedures, this means that the more active a subject is, the more successful his learning is likely to be. However, being cognitively active does not mean that the child merely manipulates a given type of material; he can be mentally active without physical manipulation, just as he can be mentally passive while actually manipulating objects. Intellectual activity is stimulated if the opportunities for acting on objects or observing other people's actions or for discussions correspond to the subject's level of development " (Inhelder, Sinclair and Bovet, 1974, p. 25).

The contrast between, on the one hand, Piaget, and on the other hand, Inhelder and co-authors, is not trivial: in a school class, the child may have many opportunities for observing other people's actions on the material, and the Inhelder-Sinclair-Bovet statement allows that such observations may be truly valuable whereas the Piaget statement quoted earlier implies that every child must perform the action himself to get any benefit from the experience.

In the training procedures reported by Inhelder, Sinclair and Bovet, the children do generally manipulate objects, with the exception of the procedure described in their Chapter 3, " From Numerical Equality to the Conservation of Matter." Even in this procedure, the child puts two sets of (modelling-clay) ' sweets ' in physical one-one correspondence (and, therefore, the experiment was omitted from our previous list) but this activity seems to be regarded as incidental and preliminary, not part of the training as such. In the rest of the training procedure, the experimenter manipulates one lot of the clay, joining and separating ' sweets ' and making elongated balls and sausages, while the other lot of clay remains unaltered; intermittently, the child is questioned about the number of sweets and about the amount of clay, comparing the two lots. If, in the early stages, the child gives incorrect answers, the experimenter returns to the initial situation (of obvious equality). The children were pre-tested, post-tested, and, after six to eight weeks, post-tested again. The tests included elementary numerical conservation, and conservation of continuous matter with clay and with liquid. Resistance to counter-suggestion was tested.

P 1504

Before training, all the children had already either acquired, or reached an intermediate level of, number conservation.

In this experiment the authors fulfilled only the first of five recommended conditions for evaluation of training experiments which they had set out. Nevertheless, they drew conclusions from the experiment, concerning the progress towards and acquisition of, conservation of matter and ' operatory schemes ' in general. Between pre-test and first post-test, some progress was recorded for 12 out of 14 children; however, of the children who were initially at the lowest level, none reached the highest level at first post-test. Between first and second post-test, performance generally was either stable or advanced (one child regressed). The authors conclude that ' genuine operatory schemes ' appear to be preceded first by an awareness of the possibility of executing a particular action which reconstitutes the initial situation. Since the children did not engage in such actions themselves, their awareness presumably came from observing the experimenter's actions with the clay.

To sum up the argument so far: Piaget (1974) and several other writers consider that an important implication of his theory is that the teacher must get the child actually to manipulate concrete material, whereas Inhelder, Sinclair and Bovet (1974), while apparently still following Piaget's theory in general, assert that it is *not* necessary for the child to manipulate the material. Furthermore there is some evidence presented by Inhelder, Sinclair and Bovet, and evidence presented by several independent researchers, which supports the Inhelder-Sinclair-Bovet assertion. Some of the experiments (Gruen, Hamel and Riksen, Brainerd) included random assignments of children to treatment groups, and were thus of ' true ' experimental design in the sense of Campbell and Stanley (1963); the other experiments did not. The experiment by Inhelder *et al.* was conducted close to the Piagetian tradition; the others were not. None of the experiments is anywhere near perfect. Some readers of the original reports will be more impressed by one experiment, others will be more impressed by another ; but a reader who refuses to give credence to any of them must be highly independent-minded. Surely the set of experiments indicates that some training procedures which do not include manipulation of objects by the child can, nevertheless, cause or provoke the development of some of those sorts of ability which comprise the theoretical stage of concrete operations.

Does this mean that part of Piaget's theory is disproved? If his hypothesis were a logical deduction from the theory, then a falsification of the hypothesis would disprove the theory. Although Piaget (1974) writes as if his hypothesis was deduced from theory, no strict logical deduction is apparent; but presumably the credibility of the theory is somewhat diminished when a hypothetical assertion derived from it by its author appears to be false, and especially when the hypothesis seems to be the theory's main practical implication.

The practical implication of the present review is not startling. Many apparently successful procedures for training concrete-operational thinking have not included (or at least have not been designed to include) manipulative activity by the child; but many *other* apparently successful training procedures *have* included such activity (Kingsley and Hall, 1967; Bearison, 1969 ; Rothenberg and Orost, 1969; Lister, 1970). Sonstroem (1966) reported that manipulation by the child was more effective than manipulation by teacher, especially when, after the child had manipulated the material, the teacher talked about it; the statistical significance is doubtful, teaching time was not controlled, and the experiment has apparently not been replicated, but the results are plausible.

No doubt, children's physical activity on material is often valuable and often under-valued; but the conclusion of the present paper is that such activity has been *over*-valued, and observation *under*-valued, by Piaget and some of his interpreters. Physical activity upon objects has a necessary place in the child's life-experience; but if we consider what the teacher should get a child to do at a particular time, then the research implies that physical activity is often *not* what is needed for intellectual development. The child observing others, in an appropriate context, need not be wasting his time. Apparently, he may reach the required understanding (he may progress to the concrete-operational 'schemes') without necessarily going through the physical motions. The extreme Piagetian insistence on physical activity has been excessive.

ACKNOWLEDGMENT—I am grateful to Mr. Richard J. Keith, for drawing my attention to the passage quoted from Piaget.

REFERENCES

ANTHONY, W. S. (1972). Psychology of education. In WOODS, R. G. (Ed.), *Education and its Disciplines.* London: Univ. of London Press.

BEARD, R. M. (1969). *An Outline of Piaget's Developmental Psychology.* London: Routledge and Kegan Paul.

BEARISON, D. J. (1969). Role of measurement operations in the acquisition of conservation. *Dev. Psychol.*, 1, 653-660.

BEILIN, H. (1965). Learning and operational convergence in logical thought development. *J. exp. Child Psychol.*, 2, 317-339.

BEILIN, H., KAGAN, J., and RABINOWITZ, R. (1966). Effects of verbal and perceptual training on water level representation. *Child Dev.* 37, 317-328.

BRAINERD, R. C. (1974). Training and transfer of transitivity, conservation, and class inclusion of length. *Child Dev.*, 45, 324-334.

BROWN, G., and DESFORGES, C. (1977). Piagetian psychology and education : time for revision. *Br. J. educ. Psychol.*, 47, 7-17.

CAMPBELL, D. T., and STANLEY, J. C. (1963). Experimental and quasi-experimental designs for research on teaching. In GAGE, N. L. (Ed.), *Handbook of Research on Teaching.* Chicago: Rand McNally.

FLAVELL, J. H. (1963). *The Developmental Psychology of Jean Piaget.* Princeton, N.J.: Van Nostrand.

GINSBURG, H., and OPPER, S. (1969). *Piaget's Theory of Intellectual Development.* Englewood Cliffs, N.J.: Prentice-Hall.

GRUEN, G. E. (1965). Experiences affecting the development of number conservation in children. *Child Dev.*, 36, 963-979.

HAMEL, B. R., and RIKSEN, B. O. M. (1973). Identity, reversibility, verbal rule instruction, and conservation. *Dev. Psychol.*, 9, 66-72.

INHELDER, B., and SINCLAIR, H. (1969). Learning cognitive structures. In MUSSEN, P., LANGER, J., and COVINGTON, M. (Ed.), *Trends and Issues in Developmental Psychology.* New York: Holt, Rinehart and Winston.

INHELDER, B., SINCLAIR, H., and BOVET, M. (1974). *Learning and the Development of Cognition.* London: Routledge and Kegan Paul.

KINGSLEY, R. C., and HALL, V. C. (1967). Training conservation through the use of learning sets. *Child Dev.*, 38, 1111-1126.

KOHNSTAMM, G. A. (1963). An evaluation of part of Piaget's theory. *Acta Psychologica*, 21, 313-356.

KUHN, D. (1974). Inducing development experimentally: comments on a research paradigm. *Dev. Psychol.*, 10, 590-600.

LISTER, C. M. (1970). The development of a concept of volume conservation in ESN children. *Br. J. educ. Psychol.*, 40, 55-64.

PIAGET, J. (1974). Foreword. In SCHWEBEL, M., and RAPH, J. (Ed.), *Piaget in the Classroom.* London: Routledge and Kegan Paul.

PIAGET, J., and INHELDER, B. (1969). Intellectual operations and their development. In FRAISSE, P., and PIAGET, J. (Ed.), *Experimental Psychology: its Scope and Method, VII. Intelligence.* London: Routledge and Kegan Paul.

ROTHENBERG, B. B., and OROST, J. H. (1969). The training of conservation of number in young children. *Child Dev.*, 40, 707-726.

SIGEL, I. E., and HOOPER, F. H. (1968). *Logical Thinking in Children.* New York: Holt, Rinehart and Winston.

SONSTROEM, A. McK. (1966). On the conservation of solids. In BRUNER, J. S., OLVER, R. R., GREENFIELD, P. M., *et al.*, *Studies in Cognitive Growth.* New York: Wiley.

VERNON, P. E. (1976). Environment and intelligence. In VARMA, V. P., and WILLIAMS, P. (Ed.), *Piaget, Psychology and Education.* London: Hodder and Stoughton.

(*Manuscript received* 21*st July*, 1976)

Br. J. educ. Psychol., 47, 25-32, 1977

SEX ROLES IN DIVERGENT THINKING

BY D. J. HARGREAVES
(*Department of Psychology, University of Leicester*)

SUMMARY. A content analysis of responses to the 'Circles' test, designed to elicit information about sex-typing, was developed on a pilot sample of 10- to 11-year-old school children. When this analysis was applied to the main sample, characteristic sex-typed response patterns were broadly replicated, as was the finding that there was no sex difference in total Fluency scores. When instructed to adopt opposite-sex response styles on a parallel form of the same test, boys and girls in the main study showed a remarkable ability to reverse their normal patterns of response. The hypothesis that children with high Fluency scores under normal conditions would be more capable of carrying out this reversal than those with low Fluency scores was not supported. It is argued that sex-typed response styles are by no means unmodifiable psychological characteristics, and that investigations of individual differences in cognition need to emphasise styles, as well as levels, of test performance.

INTRODUCTION

Bhavnani and Hutt (1972) chided research workers in the field of divergent thinking and 'creativity' for their apparent neglect of the topic of sex differences. They claim that there exists a general assumption, explicit or implicit, that sex differences in performance on divergent thinking tests do not exist. Studies which have looked at this question, whether as a primary or subsidiary aim, have produced a confusing pattern of results. Wallach and Kogan (1965), Ward (1968), and Hargreaves and Bolton (1972), for example, were unable to find any statistically significant sex differences on measures of fluency or originality. Several other workers, however, have found the well-documented verbal superiority of girls (Hutt, 1972) to be reflected in verbal fluency scores on divergent tests, although results from other measures of divergent thinking have shown little consistency. Torrance and Aliotti (1969), Olive (1972), and Bhavnani and Hutt (1972), all found that girls obtained significantly higher scores than boys on a variety of measures of verbal fluency. Whereas Torrance and Aliotti found that boys scored significantly higher on corresponding measures of originality, however, Bhavnani and Hutt found no such difference. Raina (1969), moreover, working with 180 Indian children, found that boys excelled on all the figural measures of the Torrance Tests of Creative Thinking, as well as on some of the verbal measures. Maccoby and Jacklin (1975), in their recent exhaustive review, concur with the conclusion that verbal measures of creativity reflect the verbal superiority of girls, at least from the age of about 7 onwards, whereas non-verbal measures show no clear trends.

It is clear that some of these discrepancies must arise from differences between the studies in terms of sampling of subjects, nature of tests used, conditions of administration and so on. A more fruitful approach to the problem of sex differences might be to extend the scope of enquiry beyond the straightforward 'ability testing' analyses attempted so far, and to investigate *styles*, as well as *levels*, of test performance (Hargreaves, 1974b). It may well be that sex differences in the content of divergent test responses are more clear-cut than those in overall measures of ability. If sex-typed patterns of response content can be demonstrated empirically, two further questions arise : to what extent is each sex capable of adopting the opposite-sex style when instructed to, and how might such a capability relate to divergent thinking ability?

It is well known that variations in the test situation can exert a considerable influence on divergent thinking; Vernon (1971) has reviewed some of the empirical work in this area. Hudson (1968), for example, invited sixth-form schoolboys to complete the Uses for Objects test as themselves; as Robert Higgins, the conscientious, dedicated computer engineer; and as John McMice, the uninhibited, bohemian artist. Although convergers were generally more fluent in the role of Higgins, and divergers in the role of McMice, the most striking finding was that some convergers produced many more ingenious, witty, violent and obscene responses in the role of McMice than in the other two conditions. They adopted very much more fluent, uninhibited and imaginative response styles as a result of a relatively slight change of context. The implication is that cognitive styles, at least as measured by open-ended tests, are by no means fixed psychological characteristics: we shall extend this line of investigation into the area of sex-typed cognitive styles.

Heim (1970) has drawn attention to the suggestion that creative people generally tend towards greater cross-sex identification than the average . . . "that the creative man—in whatever field—tends to have more femininity in his psychological make-up than has the less creative man, and that the creative woman tends to possess more masculine traits than does the less creative woman" (p. 139). This notion has received empirical support from the studies of Barron (1957) and Helson (1965), and was upheld by Biller *et al.* (1969) in their study of 35 kindergarten-age boys. They distinguished between mixed and consistent sex-role patterns in terms of the discrepancy between sex-role orientation (assessed by draw-a-person procedures) and sex-role preference (assessed by toy and game selections). Boys with mixed sex-role patterns scored significantly higher on divergent tests than did those with consistent sex-role patterns. We might predict that creative subjects, at least as defined in terms of divergent thinking, are more likely to be able to adopt opposite-sex response styles than are less creative subjects.

A pilot study, involving a detailed content analysis of responses to the 'Circles' test (Torrance, 1962) is described first. The main study is then designed to test three main hypotheses:

(*a*) that any characteristic sex-typed response patterns emerging from the pilot sample will be replicable on the main sample;

(*b*) that subjects will be able to adopt opposite-sex response styles on a parallel form of the 'Circles' test when instructed to do so; and

(*c*) that subjects with high Fluency scores on the 'Circles' test will be more capable of adopting opposite-sex response styles on the parallel form than will those with low Fluency scores.

PILOT STUDY—METHOD

Sample.

The sample consisted of 135 Durham school children (71 boys and 64 girls), some of whom had participated in a previous research project (Hargreaves, 1974a). The boys' ages ranged from 9–3 to 11–3 with a mean of 10–6, and those of the girls from 9–2 to 11–3 with a mean of 10–6. The results of a 't' test for uncorrelated means, which appear in Table 1, show that there is no significant difference between these two means.

Data collection.

The 'Circles' test was chosen for detailed analysis because it provides maximally 'concrete' responses which are free from verbal abstraction, and which are, therefore, likely to be easy to classify into content categories. It was administered in group form in untimed, game-like conditions as specified by Wallach and Kogan (1965). The circles were printed in sets of 20 on sheets with no official heading or written instructions; instructions were verbally conveyed by the experimenter, and there were no restrictions on the number of sheets available. Responses were scored for Fluency and Originality.

Content analysis.

A preliminary inspection of the range of responses suggested an initial five-fold classification scheme. This was applied to the responses of a random sample of 40 subjects, and subsequently modified to produce a final six-fold system as follows:

(*a*) *Life*	All representations of animate beings or their parts, e.g., faces, bodies, insects, fish.
(*b*) *Nature*	Inanimate naturally-occurring objects or their parts, e.g., planets, spiders' webs, fruit, vegetables, snowflakes.
(*c*) *Sport-Games*	Objects used in sport and games, e.g., balls, racquets, targets, yoyos.
(*d*) *Mechanical-Scientific*	Man-made instruments and machines, e.g., aeroplanes, bicycle wheels, railway trains, dials.
(*e*) *Domestic*	Objects typically found in the home, e.g., plates, tea-pots, buttons, clothes.
(*f*) *Abstract*	Signs and symbols, e.g., letters, numbers, Venn diagrams, musical notes.

This scheme was found to be easy to apply to the majority of responses, although a few (e.g., buildings, badges, ' pictures ' and scribbles using the circle as a frame) proved difficult to classify and were omitted from the analysis. The number of responses falling into each category was counted for each child.

RESULTS

Potential sex differences in age, overall Fluency and Originality scores, and raw scores in each content category were investigated by 't' tests for uncorrelated means (two-tail); the results appear in Table 1.

TABLE 1

COMPARISONS OF MEAN AGE, DIVERGENT SCORES AND RAW CONTENT CATEGORY SCORES BETWEEN THE SEXES: PILOT STUDY.

	Boys (N=71)		Girls (N=64)		
	Mean	SD	Mean	SD	t
Age (in months)	126·41	5·46	125·81	6·15	NS
Fluency	19·76	7·03	20·19	6·29	NS
Originality	10·61	9·40	9·67	7·94	NS
Life	5·46	3·30	7·33	4·25	2·85**
Nature	2·17	1·25	1·52	0·99	3·30**
Sport-Games	2·31	1·61	2·16	1·35	NS
Mechanical-Scientific	3·01	2·02	1·50	1·59	4·76***
Domestic	4·14	2·65	4·81	2·06	NS
Abstract	0·55	1·22	1·09	1·58	2·22*

*P<0·05 **P<0·01 ***P<0·001

DISCUSSION

Table 1 shows that there are no significant sex differences between the present samples of boys and girls in Fluency or Originality scores; this is consistent with the findings of previous workers in the field, who have tended to find sex differences, if at all, on verbal rather than figural tests. The significant differences between some of our content measures, therefore, are even more noteworthy in this context. Two-tail ' t ' tests were employed because the directions of these differences were not predicted; the only emergent direction which might have been anticipated is that in the ' Mechanical-Scientific ' category. This difference has the highest degree of statistical significance ($P<0{\cdot}001$) and reinforces the cultural stereotype of a dominant male interest in science and technology. That girls should produce significantly more ' Life,' 'Abstract ' and fewer ' Nature ' responses ($P<0{\cdot}01$, $0{\cdot}05$ and $0{\cdot}01$ respectively) is less obvious in terms of such stereotypes. Two predictions which we might have made using this line of explanation and which did *not* emerge from the results, moreover, are that girls should produce more ' Domestic ' and fewer ' Sport-Games ' responses than boys. Although both of these differences are in the predictable directions, neither reaches significance ($P>0{\cdot}05$).

This study provides a starting point for the further exploration of sex differences in divergent thinking styles. Our content analysis approach can evidently provide information which cannot be gained from traditional ' ability ' studies: it is developed further in the main experiment.

MAIN STUDY—METHOD

Sample.

The sample consisted of 101 children (50 boys and 51 girls) from two Leicestershire primary schools. The boys' ages ranged from 10–2 to 11–9 with a mean of 10–11, and those of the girls from 10–1 to 11–10 with a mean of 11–0. The results of a ' t ' test for uncorrelated means, which appear in Table 2, show that there is no significant difference between these two means.

TABLE 2

COMPARISONS OF MEAN AGE, DIVERGENT SCORES AND RAW CONTENT CATEGORY SCORES BETWEEN THE SEXES: MAIN STUDY.

	Boys (N=50)		Girls (N=51)		
	Mean	SD	Mean	S D	t
Age (in months)	131·32	5·18	131·55	5·56	NS
' Circles ' Test					
Fluency	29·84	11·68	30·90	21·07	NS
Life	6·98	4·35	7·69	4·71	NS
Nature	3·14	2·18	4·10	2·93	NS
Sport-Games	3·36	2·66	2·88	2·58	NS
Mechanical-Scientific	7·92	4·02	4·76	4·05	3·89***
Domestic	6·50	4·30	9·31	8·71	2·03*
Abstract	2·00	2·78	1·96	2·44	NS
' Squares ' Test					
Fluency	16·72	7·32	16·27	10·97	NS
Life	1·46	1·86	0·84	1·22	NS
Nature	0·18	0·44	0·25	0·52	NS
Sport-Games	1·60	1·76	2·92	2·57	2·98**
Mechanical-Scientific	2·44	2·35	4·20	3·71	2·80**
Domestic	10·00	5·03	7·41	5·62	2·41*
Abstract	1·04	2·00	0·65	1·18	NS

*P<0·05 **P<0·01 ***P<0·001

Procedure.

The ' Circles ' test was administered first, under the same conditions as in the pilot study. A parallel form of the test, using squares instead of circles, was then administered with the following instructions: " If you're a boy, pretend you're a girl doing these drawings; if you're a girl, pretend you're a boy doing them." After the initial amusement had died down, the children appeared to be making a genuine attempt to follow the instructions. Responses to both tests were scored for Fluency, and the content analysis was carried out in each case.

RESULTS

Potential sex differences in age, overall Fluency scores, and raw scores in each content category were investigated by ' t ' tests for uncorrelated means (two-tail); the results for both ' Circles ' and ' Squares ' appear in Table 2. ' High divergent ' and 'low divergent ' subject groups were then constructed for both sexes by dichotomising the Circles-Fluency scores at their medians. This procedure produced 26 boys and 25 girls in the ' high ' groups, and 24 boys and 26 girls in the ' low ' groups, respectively. A comparision of the *raw* content category scores of these groups on the ' Squares ' test may provide a misleading account of their relative response patterns; if Fluency scores on the two tests are positively correlated, ' high divergent ' groups will tend, on the whole, to have higher scores in all categories. To overcome this problem, scores were designed to measure opposite-sex response styles in the ' Squares ' test; these were defined as $\left(\frac{\text{Domestic}}{\text{Domestic}+\text{Mechanical-Scientific}}\right)\times 100$ for boys and $\left(\frac{\text{Mechanical-Scientific}}{\text{Domestic}+\text{Mechanical-Scientific}}\right)\times 100$ for girls. This definition is based on the finding that significantly more ' Domestic ' responses are produced by boys in the ' Squares ' test but by girls in the ' Circles ' test, and that significantly more ' Mechanical-Scientific ' responses are produced by girls in the ' Squares ' test but by boys in the ' Circles ' test (see Table 2). The ' high divergent ' and ' low divergent ' groups within each sex were compared in terms of these scores by ' t ' tests for uncorrelated means; a two-tail probability of 0·05 was taken as significant, since the directions of any emergent group differences are only predictable on empirical grounds.

There were no significant differences between the scores of the ' high divergent ' and ' low divergent ' boys (Means=80·53, 79·71 ; SD's=13·72, 22·56 respectively; t=0·15) or between those of the ' high divergent ' and ' low divergent ' girls (Means=31·82, 38·58 ; SD's=20·29, 25·99 respectively ; t=1·01).

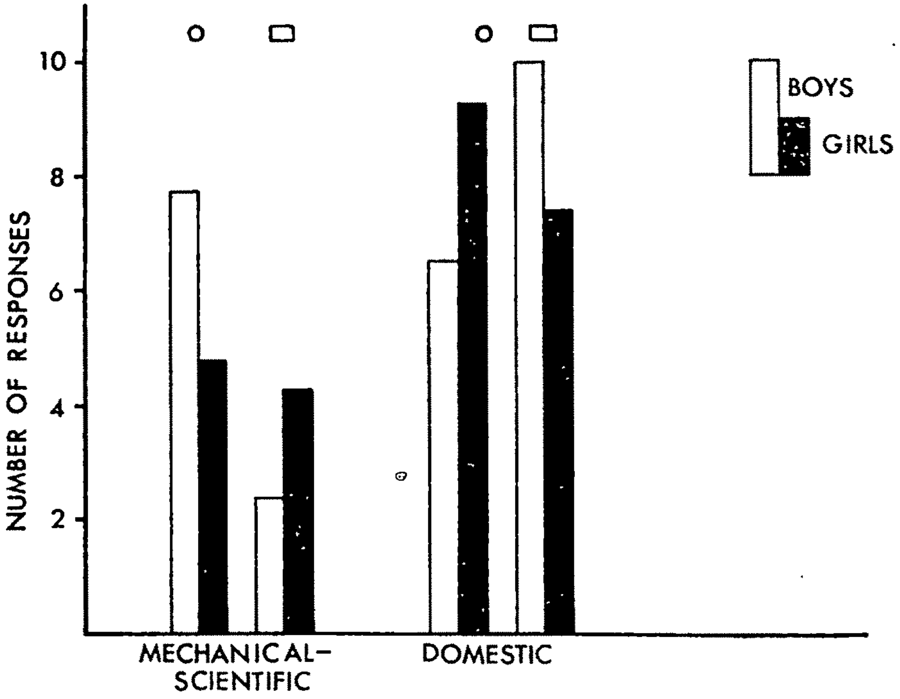
DISCUSSION

The results from the ' Circles ' test in Table 2 confirm the pilot study finding that there are no significant sex differences in overall Fluency scores ; they also provide broad support for the characteristic sex-typed response content patterns identified in the pilot study, with some minor variations. The directions of the sex differences are the same for four of the six content categories, with one particularly robust result (boys producing significantly more ' Mechanical-Scientific ' responses than girls, $P<0{\cdot}001$, in both studies). That boys produce more ' Sport-Games ' and fewer ' Domestic ' and ' Life ' responses is confirmed in the present study, though with varying levels of significance. The two categories in which the direction of sex differences is not replicated, ‘ Nature ’ and ‘Abstract,’ account for the smallest number of responses (each less than 11 per cent of the total in both studies), so are relatively unimportant.

The ‘ Squares ’ test results in Table 2 are also consistent with the finding that there are no significant sex differences in Fluency scores, though it should be noted that the total number of responses produced by both sexes is considerably lower than in the ‘ Circles ’ test. It seems reasonable to suggest that the children could not display their normal fluency under the constraint of opposite-sex responding. The results of the content analysis of the ‘ Squares ’ test are quite striking: there is a uniform reversal of the direction of each sex difference which we considered to be meaningful and consistent in the ‘ Circles ’ test. There is no such reversal in the ‘ Nature ’ and ‘ Abstract ’ categories, which we considered to be relatively unimportant. Figure 1 illustrates the two most powerful reversals; in the ‘ Mechanical-Scientific ’ and ‘ Domestic ’ categories, a statistically significant sex difference in the ‘ Circles ’ test becomes a statistically significant difference in the opposite direction in the ‘ Squares ’ test.

FIGURE 1

MEAN SCORES FOR EACH SEX IN THE ‘ MECHANICAL-SCIENTIFIC ’ and ‘ DOMESTIC ’ RESPONSE CATEGORIES.



Our analysis of the response patterns of the ‘ high divergent ’ and ‘ low divergent ’ groups produced no significant results for either sex. In the case of the girls, in fact, the results were in the opposite direction to that predicted by our hypothesis; ' low divergent ' girls produced more ' opposite-sex ' responses on the ' Squares ' test, as defined by our measure, than did ' high divergent ' girls.

CONCLUSIONS

It seems reasonable to conclude that there exist characteristic sex-typed divergent response styles, at least to the tests used in this study, and that these styles are more clear-cut than are sex differences in divergent thinking ' ability.' The major features of these styles, as found in our pilot study, were replicated in the main study, though there were some minor variations between the two samples. These results appear to reflect the developing interests of boys and girls along culturally prescribed lines. Our analysis has revealed a differentiation between the affective styles of the sexes which has been detected in the play of pre-school children (e.g., Clark *et al.*, 1969) as well as in the attitudes of adults (Allport *et al.*, 1951).

The hypothesis that boys and girls would be able to adopt opposite-sex response styles under a simple instruction to do so was strikingly confirmed in the main study; it seems clear that these styles are not fixed, unmodifiable features of psychological functioning. Children of this age can not only perceive their essential characteristics, but are evidently capable of enacting them with considerable accuracy. We need to maintain a clear distinction between the overt, behavioural manifestations of sex-typing and their covert, cognitive determinants; the results support Hudson's (1968) view that our behaviour involves a choice among different selves that exist within us.

The final hypothesis, that 'high divergent' subjects would be more capable of adopting opposite-sex response styles than ' low divergent ' subjects, was not supported by our results. It may be that the power of a simple divergent drawing test to illuminate more subtle interactive processes has reached its limits here, and that more sophisticated instruments and procedures are required. It is clear, however, that investigations of individual differences in cognitive style need to take a broader perspective than that of the traditional psychometric approach; we should emphasise the *processes*, as well as the *products*, of test performance.

ACKNOWLEDGMENTS—I should like to thank Roger Hewins, Linda Hargreaves and Raph Gillett for their help with this research.

REFERENCES

ALLPORT, G. W., VERNON, P. E., and LINDZEY, G. (1951). *The Study of Values*. Boston: Houghton Mifflin.

BARRON, F. (1957). Originality in relation to personality and intellect. *J. Personality*, 25, 730-742.

BHAVNANI, R., and HUTT, C. (1972). Divergent thinking in boys and girls. *J. child Psychol. Psychiat.*, 13, 121-127.

BILLER, H. B., SINGER, D. L., and FULLERTON, M. (1969). Sex-role development and creative potential in kindergarten-age boys. *Dev. Psychol.*, 3, 291-296.

CLARK, A. H., WYON, S. M., and RICHARDS, M. P. M. (1969). Free-play in nursery school children. *J. child Psychol. Psychiat.*, 10, 205-216.

HARGREAVES, D. J. (1974a). Situational influences on divergent thinking. *Br. J. educ. Psychol.*, 44, 84-88.

HARGREAVES, D. J. (1974b). Psychological testing: current perspectives and future developments. *Educ. Rev.*, 27, 26-33.

HARGREAVES, D. J., and BOLTON, N. (1972). Selecting creativity tests for use in research. *Br. J. Psychol.*, 63, 451-462.

HEIM, A. W. (1970). *Intelligence and Personality*. Harmondsworth: Pelican.

HELSON, R. (1965). Childhood interest clusters related to creativity in women. *J. consult. Psychol.*, 29, 352-367.

HUDSON, L. (1968). *Frames of Mind*. London: Methuen.

HUTT, C. (1972). *Males and Females*. Harmondsworth: Penguin.

MACCOBY, E. E., and JACKLIN, C. N. (1975). *The Psychology of Sex Differences*. London : Oxford University Press.

OLIVE, H. (1972). A note on sex differences in adolescents' divergent thinking. *J. Psychol.*, 82, 39-42.

RAINA, M. K. (1969). A study of sex differences in creativity in India. *J. Creat. Behav.*, 3, 111-114.

TORRANCE, E. P. (1962). *Guiding Creative Talent*. Englewood Cliffs, N.J.: Prentice-Hall.

TORRANCE, E. P., and ALIOTTI, N. C. (1969). Sex differences in levels of performance and test-retest reliability on the Torrance Tests of Creative Thinking. *J. Creat. Behav.*, 3, 52-57.

VERNON, P. E. (1971). Effects of administration and scoring on divergent thinking tests. *Br. J. educ. Psychol.*, 41, 245-257.

WALLACH, M. A., and KOGAN, N. (1965). *Modes of Thinking in Young Children*. New York: Holt, Rinehart and Winston.

WARD, W. C. (1968). Creativity in young children. *Child Dev.*, 39, 737-754.

(*Manuscript received* 23rd *April*, 1976)

Br. J. educ. Psychol., 47, 33-39, 1977

THE INFLUENCE OF CONTEXT ON REFERENTIAL DESCRIPTION IN CHILDREN

By JUDITH M. WATSON

(*University of Dundee*)*

SUMMARY. Two experiments are described which are concerned with the development of referential ability in children. In the first, children of normal and of subnormal educational attainment, aged between 4 and 12 years, were asked to describe simple actions. Referential adequacy increased markedly with mental age. The inadequacy of descriptions given by the less mature subjects was largely due to their neglect of the immediate perceived context of the referent.

The second experiment, using children in the mental age range of 5 to 7 years, provided evidence that the less mature children's descriptions were significantly less affected by the context than were those of the 7-year-olds of normal intelligence. It is suggested that this aspect of referential ability results from general cognitive decentration.

INTRODUCTION

THE experiments to be briefly described are concerned, broadly, with the changing pattern of referential description, and more narrowly, with the effect of the perceived context upon this pattern. It is claimed that the ability to make reference which is appropriate to its context logically precedes the ability to make appropriate allowance for a listener's needs.

Reference is successful when a referent (an object or an action) is identified among other possible referents. The ease of description of a referent, its codability, varies with the array from which it is to be distinguished. Thus, a zebra in a herd of zebras has low codability: one needs a full description in order to identify it. The same zebra in a herd of elephants, on the other hand, would be highly encodable, easily distinguished verbally by its name, and perceptually by its coat. The context determines the ease of reference.

The only context here considered is visual and immediate, extending not more than a few seconds in time, or a few centimetres in space. It is limited to previously or simultaneously presented non-referents. The extended context of a subject's past experiences or expectations on a larger scale is not discussed.

On the basis of adults' descriptions of simple demonstrations with perceptual variations, Osgood (1971) has described how " non-linguistic perceptual antecedents can create cognitive presuppositions in the same way that previously uttered sentences do, and that these presuppositions influence the form that descriptive sentences take."

An experiment reported by Olson (1970) illustrates the effect of such perceptually based cognitive presuppositions. Olson asked subjects to describe the location of a gold star, always placing it beneath the same white wooden circular block. The adults' descriptions of this referent block were found to vary according to the perceived non-referents present. Thus, when it was accompanied by a black circular block, subjects said " it's under the *white* one "; when presented with a white square, they'd say " it's under the *round* one," and so on. Olson's claim that the choice of words in an utterance was a function of the speaker's knowledge of the perceived alternatives was clearly upheld.

---

* Now at Psychology Department, Moray House College of Education, Edinburgh.

However, a replication of this experiment using children as subjects indicated that their descriptions only very imperfectly reflected contrasting non-referents. Both 4- and 7-year-olds gave equally long descriptions regardless of the number of alternatives. They failed to analyse the properties of non-referents to produce non-redundant identifying descriptions, and were generally dominated by salient, but not necessarily informative, perceptual features.

It has often been noted that the context is the key to the meaning of children's early language. Bloom's influential book on language acquisition (Bloom, 1970) gives several examples of the same two-word utterance being used with different meanings, in different situations.

It is also common observation that children become gradually more comprehensible in time, even without contextual support, for example, when speaking on the telephone. Many communication studies have illustrated that the ability to communicate adequately to a listener with special needs develops late. The series of studies by Flavell *et al.* (1968) concentrated on this role-taking ability. The prior ability to make adequate reference *per se* however, has been largely ignored. There are indications that children's encodings of referents are inadequate as identifying descriptions before the mental age of 7 years, approximately. Only after this ability has developed is the child able appropriately to modify or elaborate his referential descriptions to suit a listener's special needs.

In the series of experiments to be described, descriptions of referents, simple objects or actions, were elicited under varying perceptual conditions. In particular, the production of three kinds of semantic structure was investigated. All were found to be reliably produced by adults in certain perceptual situations, and to reflect their cognitive presuppositions. The first semantic structure is modification, introduced by adults when the context presents contrasting attributes or when a referent with a novel attribute is introduced. Secondly, adults make the distinction of specificity when a referent is unique or familiar: when it is one of a set or unfamiliar, adult references are non-specific. The specificity distinction is commonly made by the use of the definite article and is dependent upon perceived familiarity or uniqueness. Lastly, the spontaneous use of referential description in negative form clearly implies comparison with other perceived or expected non-referents. It is used by adults when their expectations are not fulfilled.

Selected details about two of a series of experiments are given here to illustrate the main conclusions. The perceptual conditions found to be reliably predictive of these semantic structures in adult referential description were found, with increasing mental age, to influence the production of the same structures in children's references.

EXPERIMENT ONE

The first experiment was designed to discover the approximate age at which children are capable of adequate referential descriptions, and the extent to which this depended upon particular perceptual situations.

Method.

Seven groups of 20 subjects each were used, including an adult comparison group. The subjects of normal intelligence were in groups of mean age 4, 5, 6 and 10 years. In addition, there were two groups of educationally subnormal children, of the same chronological age as the oldest normal children, and of mean mental age 7 and 8 years.

The children of normal intelligence were selected by their teachers from classes in an urban primary school, with the requirement that none should be showing undue difficulty with school work. The educationally subnormal children were in attendance at a special school, having failed to make normal educational progress at their first primary school. Their most recent Stanford-Binet IQ scores ranged from 50 to 80 points. While several of these children suffered from minor sight, speech, and emotional problems, none of the subjects was seriously handicapped in these respects.

The subjects' task was very simple. Individually they watched a series of demonstrations enacted by the experimenter, each of which consisted of the movement of a wooden square or triangle to touch another, in a tray. After each demonstration, the subject was asked simply " What happened ? " There was no recourse to a hypothetical listener. A lid was placed over the blocks following each demonstration to reduce the children's tendency to use gesticulation towards the block as part of their answer.

There were three variations in the perceptual condition of the mobile or the stationary block. The first was the condition of specificity, in which the block was either unique, or one of a similar pair. In the second condition, one of the participant blocks was coloured, and in the third, two similar blocks were used jointly in either the mobile or stationary role.

The order of presentation was varied systematically so that no two subjects in any group viewed the 24 demonstrations in the same order. There were eight demonstrations in each perceptual condition. Only subjects who had previously succeeded in naming the shapes and colours were used.

RESULTS

In analysing the recorded descriptions a major interest lay in the degree to which they were elaborated to refer to the specificity, colour, or number of the blocks, in the appropriate context.

The 4-year-olds generally described the demonstrations in a stereotyped manner, often referring only to the end product as in " it's a house," or alternatively, to the mobile block " square went." The great majority of these descriptions failed to refer to the specificity, colour, or number of the referent blocks.

TABLE 1

EXPERIMENT 1: NUMBER OF SUBJECTS INCLUDING INFORMATION APPROPRIATE TO THE PERCEPTUAL CONDITION OF THE REFERENT (N=20 in each group)

Group	Mean Age	Mean M.A.	Perceptual Condition of Referent		
			Specificity	Colour	Number
(Normal)					
1	4:2		2	1	4
2	5:2		7	7	17
3	6:4		10	7	20
4	10:2		15	16	20
5	Adult		19	20	20
1 (ESN)	10:5	6: 9	7	12	18
2 (ESN)	11:1	7:10	11	6	19

On the other hand, the descriptions given by the 10-year-old children of normal intelligence resembled those of adults, in form and in content, reliably including context dependent elaboration in all three perceptual conditions. Between the ages of 4 and 10 years an increasing number of children elaborated their descriptions according to changes in the perceptual situation of the referents, as shown in Table 1.

The educationally subnormal subjects' descriptions were less mature than might have been expected from their estimated mental age, falling on most measures between those of normal 5 and 6-year-olds. Group 1 (ESN) children in particular tended to produce apparently haphazard and inconsistent descriptions. Thus, of the 12 subjects who included some colour reference, none did so on every appropriate occasion. The descriptions given by subjects of normal intelligence were consistent, either omitting or including context appropriate elaboration in every one of a particular type of demonstration. This apparently simple task, making no great demands on vocabulary or memory, sharply differentiated educationally subnormal from normally intelligent children of the same chronological age.

Repeated measures analyses of variance were carried out separately on the normal and the educationally subnormal children's scores, for the factors of age, sex, and the perceptual condition of the referents. For the normal subjects, age was highly significant ($F=11{\cdot}02$, $P<0{\cdot}001$), as was the perceptual condition of the referents ($F=24{\cdot}02$, $P<0{\cdot}001$). For the ESN subjects, only the main factor of perceptual condition had a significant effect ($F=24{\cdot}1$, $P<0{\cdot}001$). There was a significant interaction effect between the factors of age and perceptual condition for both the normal and ESN subjects ($P<0{\cdot}05$ in each case), reflecting the fact that only the oldest and adult normal subjects elaborated their descriptions to an equal extent in the three perceptual conditions. The remaining subjects referred to number earliest and most reliably, and to colour latest.

This finding, that the younger subjects elaborated their descriptions most readily in the number situation, was thought to result from the role of the context of non-referents in the different perceptual conditions. The majority of 5-year-olds commented upon the number situations in which two shapes placed together directly participated in an action. The other two conditions, however, only were noted by the subject if he attended to the context of non-participant blocks.

The tentative suggestion arising from the first study was that at least part of the referential inadequacy of young children is due to their tendency to describe a referent in isolation from its context, in terms that do not adequately identify it from other possible referents.

EXPERIMENT TWO

The second experiment pursued the idea that younger and retarded subjects' descriptions of referents often ignore the immediate perceived context. A simple comparison was made between descriptions of referents presented alone and in a context of contrasting items, differing in shape, colour or size. It was thought that subjects whose descriptions depended upon the context would alter them in an array presentation, while others would tend to always describe the same referent in the same manner, disregarding other array elements.

Method.

There were two groups of children of normal intelligence of mean age 5 and

7 years, and two groups of educationally subnormal children of similar mental age and socio-economic status. Each group contained 16 children.

Eight stimulus items were presented in both solitary and array conditions. They consisted of simple geometrical shapes drawn on card. To facilitate group comparisons subjects were instructed to "just say one word, to show which one I'm pointing to." These instructions and two practice examples appeared to be understood by all the subjects.

Following the production task, a short comprehension task was given, in which the experimenter gave the subject a one-word identifying description, and the subject indicated the appropriate array element. The comprehension task was easily and accurately performed by all subjects.

RESULTS

Among the data from Experiment 2 are the number of identifying descriptions, and the amount of dimension switching by subjects. Those one word descriptions which successfully served to identify the referent in the array condition, like ' box,' ' square,' ' wee,' or ' blue,' in appropriate array conditions were scored as identifying descriptions.

The second measure, dimension change, was a guide to descriptive flexibility rather than accuracy. It was obtained by comparing descriptions of similar referents, in the solitary and in the array conditions, and noting those where the chosen descriptive dimension was altered.

TABLE 2

EXPERIMENT 2—NUMBER OF IDENTIFYING DESCRIPTIONS AND OF DIMENSION CHANGES IN THE ARRAY CONDITION (N=16 in each group).

Group	Mean Age	Mean M.A.	Identifying Descriptions	Dimension Change
(Normal)				
1	5:5		41	23
2	7:2		71	51
1 (ESN)	8:11	5:9	48	32
2 (ESN)	10:4	7:1	48	34

The number of identifying one word descriptions produced by the 7-year-old normal children in the array condition was significantly greater than that of the educationally subnormal children of similar mental age, and the normal 5-year-olds. (Mann-Whitney U tests between Groups 1 and 2, $U=58{\cdot}5$, $P<0{\cdot}01$: between Groups 2 and 2 (ESN), $U=12{\cdot}5$, $P<0{\cdot}025$). Similar results were obtained for the measure of dimension change. (Mann-Whitney U tests between Groups 1 and 2, $U=2$, $P<0{\cdot}001$: between Groups 2 and 2 (ESN), $U=12{\cdot}5$, $P<0{\cdot}025$.)

The prediction that older, normally intelligent, children would produce descriptions which depended upon a referent's context to a greater extent than younger subjects was upheld, and again the mentally retarded children's descriptions were less mature than might have been expected from their mental age.

DISCUSSION AND CONCLUSIONS

Throughout the series of experiments it was found, as anticipated, that referential ability increased with age. Young children are often hard to understand, and a major source of difficulty is knowing what they are talking about.

Before the age of 5 children rarely altered descriptions in accordance with contextual changes, and their descriptions were often incomplete and inappropriate. Beveridge and Tatham (1974) have concluded that a vocabulary age of at least four years is necessary for severely subnormal subjects to participate in referential communication tasks. Many children below 5 years cannot attend to the requirements of a structured test situation. They may, of course, produce adequate referential descriptions in a natural situation before this age, but generally the preferred mode of reference is gesture, or verbal indications which draw attention to the referent itself rather than to its attributes.

A growing proportion of children beyond the age of 5 gave descriptions which were obviously appropriate to the referent in question. The less mature children used names, if one was available, or noted some striking feature which might or might not differentiate the referent from its context. They also frequently used metaphorical or functional descriptions, like " it's a round sweetie," or " you can roll it." These were not useful as identifying descriptions but apparently resulted from associations in the broader context of the child's past experience with familiar objects.

The referential descriptions of 10-year-olds in Experiment One resembled those given by adults. Between the ages of 5 and 10 there was increasing differentiation of vocabulary and elaboration of content. Much of this elaboration is believed to result from the subject's awareness of the context and its effect on the adequacy of his description. The increasing frequency with mental age of the semantic dimensions of specificity, modification, and negation were direct measures of this contextual awareness.

The length of descriptions increased steadily with age in Experiment One, but this factor did not account for descriptive adequacy. The results of Experiment Two where descriptions were limited to a single word resembled those of Experiment One in the contextual neglect and metaphorical preferences of the less mature subjects.

Mental maturity was found to be a crucial factor in referential ability. Educationally subnormal children produced descriptions which resembled those of much younger children, in length, vocabulary and especially in their reliance on metaphor and disregard of context. The more mature children are apparently aware of the power of language to identify objects unambiguously. They use the stratagem of making comparisons between a referent and its surrounding context. This comparison behaviour is crucial to the development of successful verbal reference and is believed to depend on general cognitive decentration. The development of this in normally intelligent children between the ages of 5 and 8 has received ample documentation.

NOTE

Part of the material on which this paper is based was presented at the B.P.S. conference on *Language and the Social Context*, at the University of Stirling, January, 1975.

REFERENCES

BLOOM, L. (1970). *Language Development: Form and Function in Emerging Grammars.* Cambridge, Mass.: M.I.T.

BEVERIDGE, M., and TATHAM, A. (1974). *The Performance and Interaction Patterns of Mentally Handicapped Children in a Structured Communication Task.* Paper presented to the Education Section of the B.P.S., Edinburgh.

FLAVELL, J., *et al.* (1968). *The Development of Role-Taking and Communication Skills in Children.* New York: Wiley.

OLSON, D. R. (1970). Language and thought: aspects of a cognitive theory of semantics. *Psychol. Rev.*, 77, 257-273.

OSGOOD, C. E. (1971). Where do sentences come from? In STEINBERG, D. D., and JAKOBOVITZ, L. A. (Eds.), *Semantics: An Interdisciplinary Reader in Philosophy, Linguistics and Psychology.* London: Cambridge University Press.

(*Manuscript received* 14*th June*, 1976)

Br. J. educ. Psychol., 47, 40-49, 1977

REDUNDANCY IN THE CONCEPT ATTAINMENT OF ADOLESCENTS

By C. H. SINGLETON
(*Department of Psychology, University of Hull*)

SUMMARY. The production of informationally redundant choices in a selection concept attainment task was investigated in relation to age and intelligence, using 62 adolescents aged 12:7—15:6 drawn from the lower middle ability range of a comprehensive school. Subjects scoring below the 20th percentile on intelligence were found to make a significantly greater proportion of such choices when compared with subjects scoring above the 80th percentile, and the prediction of a discontinuity in the decrease in redundancy with increasing age was confirmed. Differences in redundancy due to age and intelligence were found to be of a type which is largely attributable to lack of confidence in reasoning ability, desire for verification of hypotheses, and preference for testing hypotheses directly rather than drawing inferences from information already available.

INTRODUCTION

CLARK (1971) has proposed a number of ' principles ' of concept attainment which, he suggests, might usefully be applied in the classroom learning situation. Prominent among these, by virtue of the weight of positive evidence, are the statements that as both the age and the intelligence of the child increase, so too does the ease of concept attainment. This improvement with age continues well into adolescence for subjects of average intelligence (Pishkin *et al.*, 1967) and even into early adulthood for subjects of above-average intelligence (Kates and Yudin, 1964). However, the experimental design and the performance measures typically employed in the research from which Clark's statements were derived, have generally precluded more detailed knowledge of the age- and intelligence-related processes involved. The typical experimental design is one where the experimenter controls the sequence of information presented to the child, from which the latter must learn the concept. This has been called the ' reception paradigm,' and while it may be a fairly good representation of concept attainment in the traditional classroom situation, generally it reveals relatively little about the information-seeking and information-processing strategies of the child and the information he would *like to have* in order to solve the problem. The most common measures of performance in this type of task are the number of concept instances required, the number of errors made, and the time taken by the child as he tries to attain the concept: all comparatively crude indices of cognitive ability.

Certain modifications of the reception paradigm, for example, training or instructional techniques (McKinney, 1972; Stern, 1967; Tagatz, 1967) and the use of the blank-trials procedure (Eimas, 1969, 1970; Gholson *et al.*, 1972; Ingalls and Dickerson, 1969) have facilitated the discovery of important aspects of the child's conceptual strategies which improve as a function of age and intelligence. The most notable finding has been that ' focusing ' tends to increase with age and intelligence, particularly during adolescence, supplanting other, less efficient strategies. Focusing, which has been regarded as the optimal strategy in concept attainment, involves hypothesising the relevance or irrelevance of each of the possible attributes of the concept and observing whether these hypotheses are confirmed or disconfirmed by the information presented.

A study by Yudin and Kates (1963) revealed that during adolescence there was a discontinuity in the development of concept attainment strategies. Broadly speaking, strategies which involve the modification of hypotheses on the basis of information conveyed by both present and previous instances tend to increase with age, while strategies which do not involve the systematic integration of previous information tend to decline with age. Focusing would come into the former category. It was found that both 14- and 16-year-olds were significantly superior in their strategies when compared with 12-year-olds, but there was no such differences between the two older age groups. However, Yudin (1966) discovered that this finding only held for the middle range of intelligence (IQ 96—110). In the lower range (80—95) 16-year-olds were superior to 12- and 14-year-olds, but there were no significant differences between the two younger age groups: in other words, the discontinuity had shifted downwards. In the higher range of intelligence (115—130) there was only slight evidence of discontinuity between the two younger age groups; otherwise, development was not discontinuous within this overall age range, though it could well be in a slightly younger age range. These findings were interpreted by the authors as supporting Piaget's description of the growth of formal reasoning in adolescence (Inhelder and Piaget, 1958).

An alternative, but relatively neglected approach to concept attainment which, arguably, better represents the discovery learning situation, is the ' selection paradigm ' devised by Bruner *et al.* (1956), where the subject himself selects the instances about which he requires information. The advantage of this technique is that examination of the sequence of instances chosen by the subject can reveal the pattern of decisions made, and from this his strategy may be inferred. As with the reception paradigm, the optimal strategy in this situation is focusing (Laughlin, 1973) but since the subject is in control of the sequence of information, he is also able to modify his strategy to include informationally redundant choices, which may provide additional confirmation or disconfirmation of any hypothesis about which he feels uncertain, as well as reducing the amount of information he has to extract from individual instances. Bruner *et al.* noticed that some of their subjects were prone to make redundant choices, although redundancy has not been seriously studied as a subject variable since that pioneering work. It has, however, been examined as a task variable using the reception paradigm. Bourne and Haygood (1959, 1961) found that increasing the number of redundant relevant dimensions in concept identification problems makes the task easier, and that this effect is more marked in problems with a large number of irrelevant dimensions (which would normally be rather difficult). This finding has been confirmed by Haygood and Bourne (1964), Janke (1972), Keele and Archer (1967), and Knowles *et al.* (1971), but it must be pointed out that all these investigations employed undergraduate or college-age age subjects; with younger subjects the evidence (from relatively few studies) is somewhat equivocal. Shantz (1967) found that the performance of children aged 7½—11½ on a seriation concept attainment task did not improve significantly when redundant information was introduced. On the other hand, Di Vesta and Ingersoll (1969) found that semantic redundancy had a facilitating effect upon the identification of verbal concepts by children aged 9—12, providing the subjects were oriented by the experimenter towards the pertinent cues in the task; when such orientation was not given, no significant effects of redundancy were found. From these studies it would appear that the effects of stimulus redundancy on the concept attainment of children are highly task specific.

Bruner *et al.* suggested that redundant choices were due largely to the strain upon memory that strategies other than focusing tend to create. As focusing has been shown to increase with age and intelligence, one might therefore predict that redundancy should exhibit a corresponding decrease. Such a prediction would be consistent with Wohlwill's (1962) theoretical generalisation that conceptual development involves a decrease in the amount of redundant information required. Moreover, on the basis of the findings of Yudin and Kates (1963) and Yudin (1966) one would also predict a *discontinuity* in the decrease in redundant choices during adolescence, because strategies which do not involve the systematic integration of previous information have been shown to decline discontinuously, and such strategies will inevitably result in redundant choices. The point in development at which this discontinuity should occur may be predicted from the intelligence range of the sample: the lower the intelligence of the sample, the later in adolescence will the discontinuity be. The present study was designed to test these predictions.

METHOD

Sample.

All 173 pupils aged from 12:7 to 15:6 attending an urban comprehensive school, except those in the 'A' stream, were given the AH4 group test of intelligence. The median age of the sample was 14:2; the mean AH4 score was 57·50 (S.D.=15·52). Fifteen experimental subjects were chosen at random from each of three age ranges: I. 12:7—13:6, II. 13:7—14:6, III. 14:7—15:6; the mean AH4 scores of these three groups of experimental subjects were not significantly different from those of the original samples from which they were drawn. In addition, within the second age range, all pupils with AH4 scores above the 80th percentile for that group, and all pupils with AH4 scores below the 20th percentile for that group, were also selected for experimental testing. The former category (termed the ' high ' intelligence group) had a mean AH4 score of 77·67 (S.D.=6·50), while the latter category (termed the ' low ' intelligence group) had a mean AH4 score of 39·85 (S.D.=5·84).

It must be emphasised that the use of the terms ' high ' and ' low ' with regard to the intelligence groups is purely relative to this particular sample. Since the mean AH4 score of the whole sample was slightly below that for 565 14-year-old secondary modern school children (Mean=59·51, S.D.=19·50) cited in the AH4 test manual (Heim, 1967), and the standard deviation substantially smaller, it is evident that this sample should be regarded as being drawn from the lower middle range of the intelligence spectrum. The mean AH4 score of the ' high ' group is below that for 533 14-year-old grammar school children (Mean=85·26, S.D.=11·21) also cited in the test manual, so that, in this case, ' high ' should certainly not be taken to mean ' gifted.'

The sizes of these experimental groups were determined by the duration for which the school facilities were available; this was outside the control of the author. A better experimental design would have been to establish ' high,' ' medium ' and ' low ' intelligence groups for each of the three age groups, but this was not possible in this case.

The experimental testing had to be conducted out of school hours on a voluntary basis, and a few of the pupils selected failed to attend; these were subsequently dropped from the study but no attempt was made to replace them as this would have meant interfering with the basis upon which the experimental groups had already been established.

Materials.

The stimulus materials for the concept attainment problems consisted of coloured geometrical shapes which varied in three dimensions: shape, colour, and number. Each dimension had four possible values, hence the total number of possible instances was 64. Each instance was drawn on white card measuring 4 inches by 3 inches, and the total array was presented to the subject in an ordered display. Underneath each card was a button which, when pressed, caused one of two lamps to illuminate; these provided the subject with information on the status (positive or negative) of that particular instance. An indicator light by each instance obviated the necessity of the subject having to remember his current choice; this and the status lamp remained illuminated until the subject made his next choice. An array of switches on the rear of the apparatus enabled the experimenter to change the concept to be attained for each new problem.

Procedure.

The experimental testing took place individually in a quiet room with no distractions. Each subject was first shown the apparatus and his attention was drawn to the various dimensions and dimension values of the stimulus array. It was explained that in this situation a ‘ concept ’ referred to a set of instances with a feature or features in common (only conjunctive concepts were used throughout the experiment), and the subject was encouraged to point out common features shared by groups of cards in the stimulus array. The mode of operation of the apparatus was then demonstrated and five concept attainment problems were given for practice, by the end of which all subjects had successfully mastered the basic requirements of the task.

Following the practice session, each subject was given 10 concept attainment problems, of which half involved one-dimensional concepts, and the remainder two-dimensional concepts. Each problem commenced with the experimenter specifying a positive instance (the ‘ exemplar ’) after which the subject proceeded to choose further instances about which he desired information, by pressing the button under the appropriate card. This continued until the subject announced the correct concept or until he had made six choices, which had been set as the limit. The subject was informed of this in advance, and was told that for all problems it was possible to attain the correct concept within the prescribed limit. Over all subjects the mean number of problems unsolved within the limit was 0·69. As the purpose of this experiment was to examine successful, rather than unsuccessful concept attainment, whenever a subject failed to attain the correct concept within the limit he was given a substitute problem, so that data were obtained on 10 successfully completed problems for each subject. There were no significant differences between the experimental groups in the number of unsolved problems, and the incidence of such failures was too small to permit further analysis.

The sequence of instances chosen by each subject for each problem was recorded, as was the time taken from initial presentation of the exemplar to announcement of the correct concept. Each separate sequence was analysed by comparing individual choices with the examplar, with all previous choices, and with all previous pairs of choices. There were four possible outcomes of this comparison. If a choice furnished new information about the concept it was classed as an *informational* choice ; if it furnished information which was relevant to the concept but which could have been obtained from previous choice(s), it was classed as a *redundant* choice; if it was identical with any previous choice or with the examplar it was classed as a *repeat* choice; and if it had no feature in common with the exemplar it was classed as a *non-informational* choice. The last two categories were not particularly germane to the present inquiry, and as choices falling into these were infrequent, they were combined into a category which will be referred to simply as ‘ other ’ choices.

The data for each subject thus comprised the total number of choices over 10 successfully completed problems, the mean time per problem, and the proportion of choices falling into the informational, redundant, and ‘ other ’ categories, each expressed as a percentage of the total number of choices.

RESULTS

Descriptive statistics of the concept attainment performance of the three age groups are given in Table 1.

TABLE 1

MEANS AND STANDARD DEVIATIONS OF THE DEPENDENT VARIABLES FOR THE THREE AGE GROUPS.

	Age Groups					
	12:7 — 13:6 Mean	SD	13:7 — 14:6 Mean	SD	14:7 — 15:6 Mean	SD
Time per problem (secs.)	68·45	35·81	63·28	25·17	36·43	12·42
Total choices	42·83	7·17	47·14	6·75	44·08	6·84
Percentage of informational choices	56·59	12·95	53·67	9·02	62·38	7·57
Percentages of redundant choices	40·11	14·03	42·48	9·36	34·77	8·91
Percentage of ‘ other ’ choices	3·30	2·12	3·85	3·75	2·85	3·79
N	12		14		12	

All these data were analysed using the Mann-Whitney ‘ U ’ Test (Siegel, 1956). Where differences were in the predicted direction, one-tailed tests were employed; elsewhere, two-tailed tests were used. No significant differences were found between age group I and age group II. Age group III were found to have taken less time per problem ($P < {\cdot}001$), made proportionately more informational choices ($P < {\cdot}01$) and fewer redundant choices ($P < {\cdot}025$) when compared with age group II. There were no significant differences between age groups II and III in the total number of choices made, nor in the proportion of ‘ other ’ choices.

TABLE 2

MEANS AND STANDARD DEVIATIONS OF THE DEPENDENT VARIABLES FOR THE ‘ HIGH ’ AND ‘ LOW ’ INTELLIGENCE GROUPS.

	‘ High ’ Intelligence Group		‘ Low ’ Intelligence Group	
	Mean	SD	Mean	SD
Time per problem (secs.)	39·26	14·09	86·28	52·78
Total number of choices	42·47	6·51	46·62	8·60
Percentage of informational choices	61·43	8·23	49·69	8·75
Percentage of redundant choices	34·74	7·89	46·82	9·43
Percentage of ‘ other ’ choices	3·83	3·24	3·49	2·85
N	15		13	

Descriptive statistics of the concept attainment performance of the ‘ high ’ and ‘ low ’ intelligence groups are given in Table 2. The ‘ low ’ intelligence group was found to have taken more time per problem ($P<\cdot001$), made proportionately fewer informational choices ($P<\cdot01$), and more redundant choices ($P<\cdot001$), when compared with the ‘ high ’ intelligence group. There were no significant differences between the groups in the total number of choices made, nor in the proportion of ‘ other ’ choices.

The redundant choices made by each subject were further analysed into two categories: those which were redundant with respect to the immediately preceding choice (type A), and those which were not redundant with respect to the immediately preceding choice, but which were redundant with respect to choices further back in the sequence (type B). It was observed that a large proportion of type A redundant choices were of one particular form, which involved first making an informational choice by changing the value of only one attribute of the exemplar (focusing) and then following this immediately with a choice in which the same attribute is changed to a third value, the values of other attributes being unaltered. The latter choice is clearly redundant as it cannot yield any new information. For example, if the exemplar is ‘ three blue squares,’ the subject might choose ‘ three blue circles.’ which can provide information on the relevance of ‘ squares ’ in the concept, but if this is followed by a choice of ‘ three blue triangles ’ or ‘ three blue crosses ’ then such a choice would be redundant. Type A redundant choices of this particular form will be referred to as type A1, and all other type A redundant choices as type A2. The results of this categorisation of redundant choices is given in Table 3.

TABLE 3

MEANS AND STANDARD DEVIATIONS OF THE NUMBER OF REDUNDANT CHOICES OF FOUR BASIC TYPES MADE BY THE EXPERIMENTAL GROUPS.

Redundancy Type		Intelligence		Age Groups		
		‘ High ’ Group	‘ Low ’ Group	12:7—13:6	13:7—14:6	14:7—15:6
A	M..........	11·06	18·10	13·93	16·43	10·75
	SD	5·62	7·19	7·20	7·13	5·12
B	M..........	4·20	4·23	4·67	4·07	5·08
	SD	3·21	2·68	3·06	2·06	4·18
A1	M..........	9·27	14·79	11·60	13·36	8·61
	SD	5·38	5·57	5·01	5·46	5·38
A2	M..........	1·79	3·31	2·33	3·07	2·14
	SD	1·71	3·75	2·61	3·67	1·21

When compared with the ‘ high ’ intelligence group, the ‘ low ’ intelligence group was found to have made significantly more redundant choices of type A ($P<\cdot01$) and type A1 ($P<\cdot025$); no significant differences between these two groups were found in the number of redundant choices of types B and A2.

Comparing the three age groups, no signficant differences on any of the four types of redundant choice were found between the 12:7—13:6 age group and the 13:7—14:6 age group, while the 14:7—15:6 age group were found to have made significantly fewer redundant choices of types A and A1 ($P < \cdot 025$ in both cases). The oldest age group was not significantly different from the two younger age groups on number of redundant choices of types B and A2.

DISCUSSION

Although the sample sizes in this study were relatively small, the results indicate that redundant choices are a major feature of the performance of adolescents in this type of concept attainment task. Redundant choices ranged from 34 per cent to 46 per cent of all choices, and were found to be a significant function of both age and intelligence, thus confirming the predictions put forward in the introduction. Moreover, the findings of no significant difference in redundant choices between the 12:7—13:6 age group and the 13:7—14:6 age group, but a significantly smaller proportion of redundant choices in the 14:7—15:6 age group, confirms the prediction of a discontinuity in development during adolescence. This discontinuity was also found in the proportion of informational choices and in the mean time per problem. It has already been pointed out that the sample from which these experimental subjects were drawn fell into the lower middle range of intelligence, and on this basis one would predict significant improvement in information processing ability and conceptual strategy later, rather than earlier in adolescence (Yudin, 1966). The results of this investigation are congruent with such a prediction.

The important question is: why do redundant choices occur? Type B redundant choices could well be a function of memory, the result of the subject either forgetting information derived from previous instances or failing to store information efficiently in memory in the first place. Alternatively, the subject, realising the limitations of his memory, may not attempt to store all the information provided by each instance. It is difficult to assign relative importance to these potential causative factors. On the basis of evidence drawn from developmental studies of concept learning employing the reception paradigm 'blank trials' procedure, Eimas (1970) has argued that a major portion of the difficulty that younger and less intelligent children have in processing information pertains to a deficiency in memory rather than an inability to use the logical rules required for the task. Eimas maintains that this deficiency is found in both the input and the retention of information in memory. However, the literature on memory development is consistent in suggesting that forgetting rate decreases with neither age nor intelligence; the superiority of older and more intelligent children is apparently due to differences in the strategies of acquisition and retrieval of information. The most notable of these mnemonic strategies is the ability to organise information into conceptual categories, which increases with both age and intelligence (Belmont and Butterfield, 1969; Flavell, 1970; Mandler and Stephens, 1967; Neimark *et al.*, 1971). These studies support the view that the growth of memory is primarily a result of increasingly sophisticated conceptual strategies (Neimark, 1975), and they suggest that if type B redundant choices are a result of memory deficiency, this deficiency probably lies in the techniques the subject has for categorising the information and organising it in memory rather than in retention of information.

However, it is equally likely that type B redundant choices could result from the subject's inability or unwillingness to recognise all the information which each instance is capable of supplying. This view would be consistent with Wason's findings (Wason, 1960, 1968; Wason and Johnson-Laird, 1972) that even intelligent adults tend to operate a ' verification ' strategy in conceptual tasks. This strategy involves generating a single hypothesis at a time, seeking to verify rather than falsify that hypothesis, ignoring most, if not all of the information which is not strictly relevant to the current hypothesis, and generating a considerable amount of redundant information in ' support ' of that hypothesis. Such an explanation can also be applied to type A redundant choices: type A1 would reflect the subject's strong desire for verification, while type A2 could arise when the subject, changing his hypothesis on the basis of disconfirming evidence derived from the current choice, fails to notice that the new hypothesis is also disconfirmed by the current choice. His next choice to test this new hypothesis will obviously be redundant.

An alternative explanation for type A2 redundant choices could lie in the subject's desire to make a 'direct test' of an hypothesis rather than relying upon inference which involves transformation of information. Bruner *et al.* (1956) observed in some of their subjects a reluctance to make 'in the head' transformations, and a preference for the ' direct test ' even though it is strictly redundant to do so. For example, if the exemplar is ' four red crosses ' the subject might hypothesise that the concept is ' crosses ' and choose the instance ' two yellow crosses ' in an attempt to confirm that hypothesis; if the information indicates that this instance is a member of the concept in question, then all subsequent choices must be redundant. However, rather than inferring that all other hypotheses have been eliminated, the subject may prefer to test each one directly. Studies by Donaldson (1959) and by Wason and his colleagues (Wason, 1959; Wason and Johnson-Laird, 1972) on the difficulties that children and adults have in processing negative information suggest that the desire for the direct test will be particularly strong when the subject has to infer what hypotheses have been ruled out by any given instance. The evidence from the present study is consistent with this view. It is not clear, however, whether distrust of negative information arises because the subject is alive to the possibility of making errors in transforming such information to the positive (Bruner *et al.*, 1956), or because the subject feels that negative information is ' not such good currency ' as positive information (Donaldson, 1959), or because negative information is neglected in everyday life and in education and consequently the subject lacks experience with it (Wason, 1959). It is probable that in concept attainment all these three reasons are relevant.

Over all subjects, only 27 per cent of redundant choices were of type B. The remaining 73 per cent were of type A. The analysis indicates that differences in redundancy due to age and intelligence are almost entirely confined to choices of type A, and in particular, to those of type A1. Clearly memory cannot be a factor here, as information on both the content and the status of the current choice was constantly displayed to the subject until he actually pressed the button to make his next choice. Nor can the explanation lie in the subject changing his hypothesis and failing to notice that the new hypothesis is disconfirmed by the current choice (which could apply in the case of type A2 redundancy), since with type A1 redundancy there is no evidence whatsoever of subjects altering their hypotheses. Three possible explanations would seem to remain. One is the preference for the direct test, which appears a particularly likely candidate where the preceding choice was negative. However, in no fewer than 87 per cent of cases, type A1 redundant choices were preceded by positive choices. In such cases, the subject gives every appearance of following Wason's ' verification ' strategy, although even here the desire for the direct test cannot be completely ruled out, nor can the third possibility, which is that the subject simply lacks confidence in his reasoning ability, so that he has to make ' doubly sure ' or even ' trebly sure ' that he has drawn the correct inference. This lack of confidence certainly does not appear in the protocols of Wason's subjects (Wason, 1960, 1968), who tended to act as though strongly convinced that their current hypothesis was the correct one. On the other hand, Wason's subjects were mostly undergraduates, who might be expected to be fairly confident of their reasoning ability. Such confidence seems unlikely in the adolescents of average or below average ability who were employed in the present study.

The findings of this study suggest that many adolescents, particularly younger or less able ones, will tend to seek redundant information in concept attainment tasks. However, it is not certain that supplying such information always has the desired effect of facilitating reasoning, since attempts to do this experimentally with children have produced equivocal results (Shantz, 1967; Di Vesta and Ingersoll, 1969). In the present investigation, differences in redundancy due to age and intelligence were found to be of a type which is largely attributable to lack of confidence in reasoning ability, desire for verification of hypotheses, and preference for testing hypotheses directly rather than drawing inferences from information already available. These three factors would seem reducible to the single factor of lack of relevant training and experience in reasoning, which, potentially, could be remedied.

REFERENCES

BELMONT, J. M., and BUTTERFIELD, E. C. (1969). The relations of short-term memory to development and intelligence. In LIPSITT, L.P., and REESE, H. W. (Eds.), *Advances in Child Development and Behaviour* (Vol. 4). New York: Academic Press.

BOURNE, L. E., and HAYGOOD, R. C. (1959). The role of stimulus redundancy in the identification of concepts. *J. exp. Psychol.*, 58, 232-238.

BOURNE, L. E., and HAYGOOD, R. C. (1961). Effects of redundant relevant information upon the identification of concepts. *J. exp. Psychol.*, 61, 259-260.

BRUNER, J. S., GOODNOW, J. J., and AUSTIN, G. A. (1956). *A Study of Thinking*. New York: Wiley.

CLARK, D. C. (1971). Teaching concepts in the classroom. *J. educ. Psychol.*, 62, 253-278.

DI VESTA, F. J., and INGERSOLL, G. M. (1969). Effect of semantic redundancy on children's identification of verbal concepts. *J. exp. Psychol.*, 82, 360-365.

DONALDSON, M. (1959). Positive and negative information in matching problems. *Br. J. Psychol.*, 50, 253-262.

EIMAS, P. D. (1969). A developmental study of hypothesis behaviour and focusing. *J. exp. Child Psychol.*, 8, 160-172.

EIMAS, P. D. (1970). Effects of memory aids on hypothesis behaviour and focusing in young children and adults. *J. exp. Child Psychol.*, 10, 319-336.

FLAVELL, J. H. (1970). Developmental studies of mediated memory. In LIPSITT, L. P., and REESE, H. W. (Eds.), *Advances in Child Development and Behaviour*. Vol. 5. New York: Academic Press.

GHOLSON, B., LEVINE, M., and PHILLIPS, S. (1972). Hypotheses, strategies, and stereotypes in discrimination learning. *J. exp. Child Psychol.*, 13, 423-446.

HAYGOOD, R. C., and BOURNE, L. E. (1964). Forms of relevant stimulus redundancy in concept identification. *J. exp. Psychol.*, 67, 392-397.

HEIM, A. W. (1967). *AH4 Group Test of Intelligence Manual* (Revised edition). Slough, Bucks: National Foundation for Educational Research.

INGALLS, R. P., and DICKERSON, D. J. (1969). Development of hypothesis behaviour in human concept identification. *Devel. Psychol.*, 1, 707-716.

INHELDER, B., and PIAGET, J. (1958). *The Growth of Logical Thinking from Childhood to Adolescence*. London: Routledge and Kegan Paul.

JANKE, M. (1972). Effect of stimulus and instructional variables in an ambiguous concept attainment task. *J. exp. Psychol.*, 93, 21-29.

KATES, S. L., and YUDIN, L. W. (1964). Concept attainment and memory. *J. educ. Psychol.*, 55, 103-109.

KEELE, S. W., and ARCHER, E. J. (1967). A comparison of two types of information in concept identification. *J. verb. learn. verb. Behav.*, 6, 185-192.

KNOWLES, B. A., HAMMOND, K. R., STEWART, T. R., and SUMMERS, D. A. (1971). Positive and negative redundancy in multiple cue probability tasks. *J. exp. Psychol.*, 90, 158-159.

LAUGHLIN, P. R. (1973). Selection strategies in concept attainment. In SOLSO, R. L. (Ed.), *Contemporary Issues in Cognitive Psychology*: *The Loyola Symposium*. Washington, D.C.: V. H. Winston.

McKINNEY, J. D. (1972). Developmental study of the acquisition and utilisation of conceptual strategies. *J. educ. Psychol.*, 63, 22-31.

MANDLER, G., and STEPHENS, D. (1967). The development of free and constrained conceptualisation and subsequent verbal memory. *J. exp. Child Psychol.*, 5, 86-93.

NEIMARK, E. D. (1975). Intellectual development during adolescence. In HOROWITZ, F. D. (Ed.), *Review of Child Development Research* (Vol. 4). Chicago and London: University of Chicago Press.

NEIMARK, E. D., SLOTNICK, N. S., and ULRICH, T. (1971). Development of memorisation strategies. *Devel. Psychol.*, 5, 427-432.

PISHKIN, V., WOLFGANG, A., and RASMUSSEN, E. (1967). Age, sex, amount, and type of memory information in concept learning. *J. exp. Psychol.*, 73, 121-124.

SHANTZ, C. U. (1967). Effects of redundant and irrelevant information on children's seriation ability. *J. exp. Child Psychol.*, 5, 208-222.

SIEGEL, S. (1956). *Nonparametric Statistics for the Behavioural Sciences*. New York: McGraw-Hill.

STERN, C. (1967). Acquisition of problem-solving strategies in young children and its relation to verbalisation. *J. educ. Psychol.*, 58, 245-252.

TAGATZ, G. E. (1967). Effects of strategy, sex and age on conceptual behaviour of elementary school children. *J. educ. Psychol.*, 58, 103-109.

WASON, P. C. (1959). The processing of positive and negative information. *Quart. J. exp. Psychol.*, 11, 92-107.

WASON, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quart. J. exp. Psychol.*, 12, 129-140.

WASON, P. C. (1968). 'On the failure to eliminate hypotheses . . .'—a second look. In WASON, P. C., and JOHNSON-LAIRD, P. N. (Eds.), *Thinking and Reasoning*. Harmondsworth: Penguin, 165-174.

WASON, P. C., and JOHNSON-LAIRD, P. N. (1972). *Psychology of Reasoning*: *Structure and Content*. London: Batsford.

WOHLWILL, J. F. (1962). From perception to inference: a dimension of cognitive development. In KESSEN, W., and KUHLMAN, C. (Eds.), Thought in the Young Child. *Monogr. soc. res. Child Devel.*, 27, No. 2.

YUDIN, L. W. (1966). Formal thought in adolescence as a function of intelligence. *Child Dev.*, 37, 697-708.

YUDIN, L. W., and KATES, S. L. (1963). Concept attainment and adolescent development. *J. educ. Psychol.*, 54, 177-182.

(*Manuscript received 6th April*, 1976)

Br. J. educ. Pychol., 47, 50-59, 1977

THE MEASUREMENT OF SPEED OF PROBLEM SOLVING AND ITS RELATION TO CHILDREN'S AGE AND ABILITY

By C. D. ELLIOTT and D. J. MURRAY
(*University of Manchester*)

SUMMARY. The times children take to reach correct solutions to seven easy Block Design items from the try-out version of the British Intelligence Scale were analysed in terms of two regression models. The times were obtained from 395 children aged between 10 and 13 years, each of whom had a score of four or more on the full 16-item test. The simpler of the two models was found to be the most satisfactory, the model specifying that the time to correct solution of a person on a test item is a proportional function of the difficulty of the item and a single parameter describing the speed of the person. Individual speed parameters were estimated for each child and were analysed in terms of an age × ability analysis of variance. This analysis indicated highly significant main effects, speed increasing with both age and ability. There was no interaction between age and ability. Results were related to previous findings and supported a view of mental speed as a major component of ability.

INTRODUCTION

As long ago as 1926, Peak and Boring expressed the opinion that speed should be the prime factor to be studied if we are ultimately to come to an understanding of the nature of intelligence. Perhaps because of the enormous success and practical utility of the Binet, Wechsler and group tests of intelligence, this important observation has only been followed up in fairly desultory manner (see, for example, Thorndike, 1926; McFarland, 1930; Sutherland, 1934; Studman, 1935; Thurstone, 1938; Slater, 1938, 1944; Baxter, 1941; Davidson and Carroll, 1945; and Tate, 1950). The relative lack of interest in speed as a major intellectual factor has certainly not been due to its lack of relevance to cognitive performance. Indeed, virtually all group tests of ability impose time limits, faster individuals tending to get higher ability scores. In the field of individual tests, time limits have been extensively used, time credits being given to scores on a number of Wechsler performance tests, for example.

Giving individuals additional credit for rapid performance makes the test one of cognitive 'power,' this term being commonly used in science to denote rate of work. Tests of 'power' should be distinguished, but seldom are, from tests of 'level' which are measures of the maximum level of difficulty at which a person can function without time pressure.

Conventional time-limit tests do not readily lend themselves to the measurement of speed of problem solving. Total scores are, as Furneaux (1960) has suggested, determined through an interaction of three components: speed, accuracy and continuance (or persistence). The same score achieved by different persons may be due to the operation of different levels of these characteristics in each person. The only way in which the relative levels of the parameters may be estimated is by going back to item data and analysing the speed of response together with information as to whether the solution was correct, incorrect, or abandoned. Rasch (1960), White (1973) and Van der Ven (1971, 1972, 1973a, 1973b, 1974) have all described mathematical models for analysing performance on time-limit tests incorporating speed and accuracy parameters, together, in the case of the White model, with persistence. Such procedures are computationally cumbersome, and while it may in a sense be elegant to estimate two or three parameters from one test performance, one's measurement may be more effective if done with tests designed specifically for a given purpose, as Furneaux (1956) has attempted with his separate tests of speed and level.

One of the major difficulties in estimating speed from cognitive test performance is in obtaining adequate measures of item difficulty. Item difficulty, in the conventional sense of the proportion of people passing an item is closely related to the average length of time required for solution, although, as Furneaux (1960) has observed, there is by no means a one-to-one relationship between the two variables. If we wish to establish an item difficulty scale for the measurement of speed, it is clearly preferable to express the item difficulties in the same sort of metric, namely, a time-difficulty scale. Such a scale was employed by Furneaux (1960) and is also suggested by Rasch (1960) and by Van der Ven (1973).

If a set of problems are administered to a sample of persons, the mean solution time for each item obtained by the sample will be the best estimator of the time difficulty of the item. Note here, however, that the items should be so easy (in conventional terms) that they should be passed by nearly everybody in the sample. If this condition is not met, the estimate of the time difficulty of items will become progressively more biased as difficulty increases—for two reasons. Firstly, only times to *correct* solution should be used for time difficulty estimation. Times to incorrect solution may be biased by a number of interfering factors, such as guessing or lack of persistence. Secondly, if solely times to correct solution are analysed, these would be obtained only from the more intelligent subjects who would be expected to be faster than less intelligent subjects. Hence the time difficulties would be underestimated. This would lead for many individuals, to a curvilinear plot between their correct solution times and the time difficulties of the items.

TWO MODELS FOR THE ANALYSIS OF SPEED

Having briefly outlined some of the problems likely to be encountered in speed measurement, two models may now be outlined which could be used in such an analysis. Both models presuppose a linear relationship between a person's times to solution and the time difficulties of the problems in the test. Model I is very much an empirical model whereas Model II, although simpler, has a sounder theoretical basis.

Model I.

On the assumption that a linear relationship exists between difficulty and an individual's solution times, a two-parameter regression model may be postulated, such that an individual person's time to solution of any item (T_{pi}) is a function of two constants peculiar to the individual (a_p and b_p), together with the time difficulty of the item (D_i):

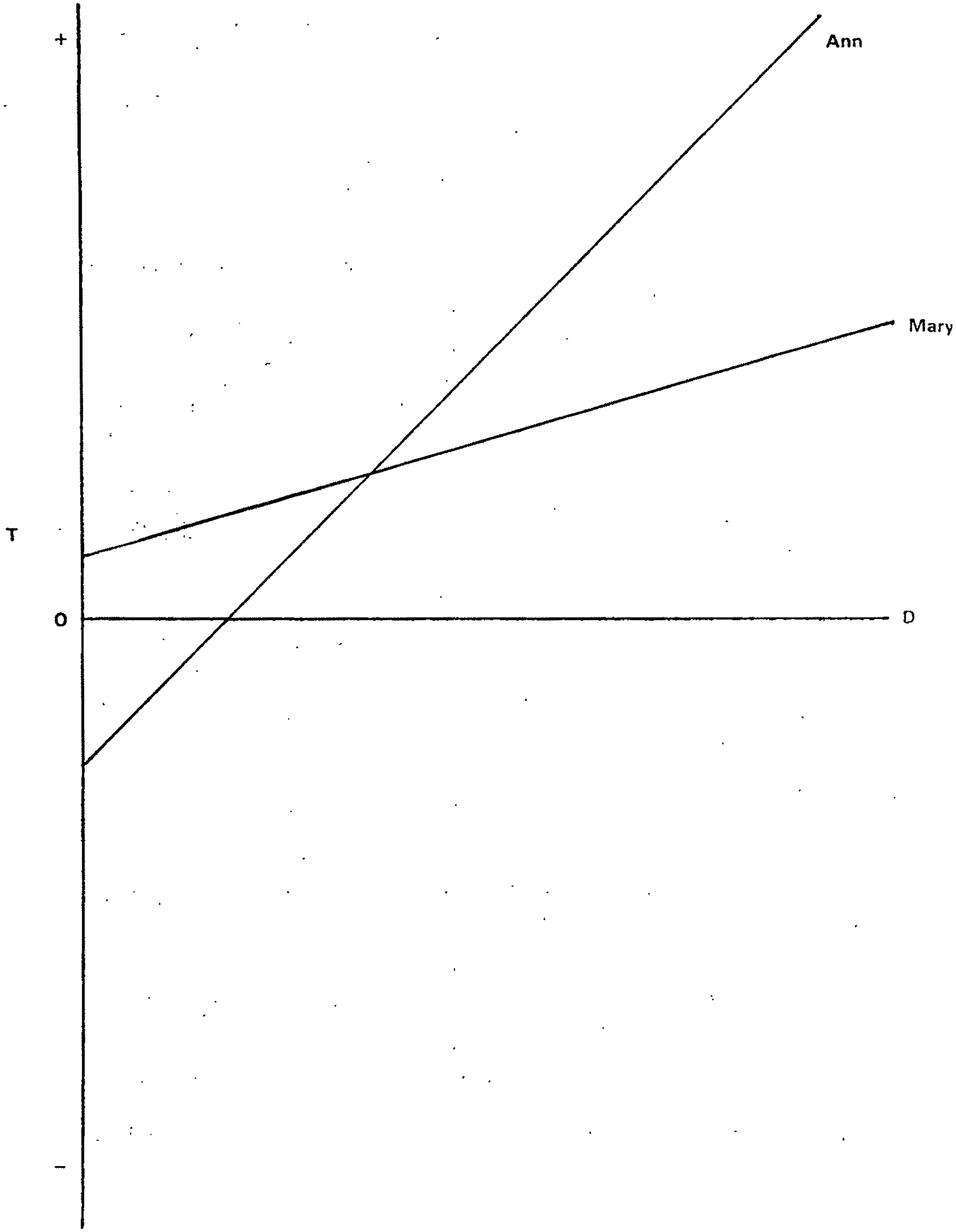
$$T_{pi}=a_p+b_p\ D_i$$

This is, of course, the standard type of regression equation derivable from a product-moment correlation (e.g., Guildford, 1956).

The regression coefficient b_p represents the slope of the relationship between T_p and D. Slower persons will have steep slopes in comparison with faster persons. Figure 1 illustrates the regression lines of two hypothetical individuals. Ann is slower than Mary, in that for a given increase in difficulty, her solution times increase at a faster rate.

FIGURE 1

Two-parameter Model $T_{pi} = a_p + b_p D_i$



The regression constant a_p represents the person's intercept on the ordinate T for a value of $D_i=0$. There is no theoretical basis for this constant. It may be argued, however, that incorporating this second parameter into the model will result in a better fit for any data which we wish to analyse. If it were necessary to describe a person's speed of problem solving in terms of at least two parameters, this might indicate that the underlying determinants of speed are relatively complex.

Model II.

In the second model the constant a_p has been dropped for a good theoretical reason. As a_p represents the intercept on the ordinate when $D_i=0$, it is inconceivable that any individuals could have T values of anything other than zero. D_i, it will be recalled, are determined on the basis of the sample mean times to correct solution. In the hypothetical case of an item being of zero difficulty, no person in the sample would take any time at all to get it right, let alone the ludicrous possibility of solving it in negative time. The second theoretical model thus states that an individual's time to solution of an item is a function of the difficulty of the item and just one constant peculiar to that individual:

$$T_{pi}=b_p \ D_i \,.$$

This model states that individuals differ only in their speed slopes, as shown by the solid lines in Figure 2. If the model were to hold, the calculation of slopes in the context of a practical test administration would be relatively easy, for, in a test of n items, the best estimate of an individual's slope would be:

$$b_p = \frac{\sum_{i=1}^{n} \frac{T_{pi}}{D_i}}{n}$$

However, as the distribution of an individual's response times may be positively skewed, the slope parameter estimated from all available items would be greater than that derived from his most characteristic or modal performance. This tendency might be reduced by dropping from the estimation a number of response times which are markedly long in relation to item difficulty.

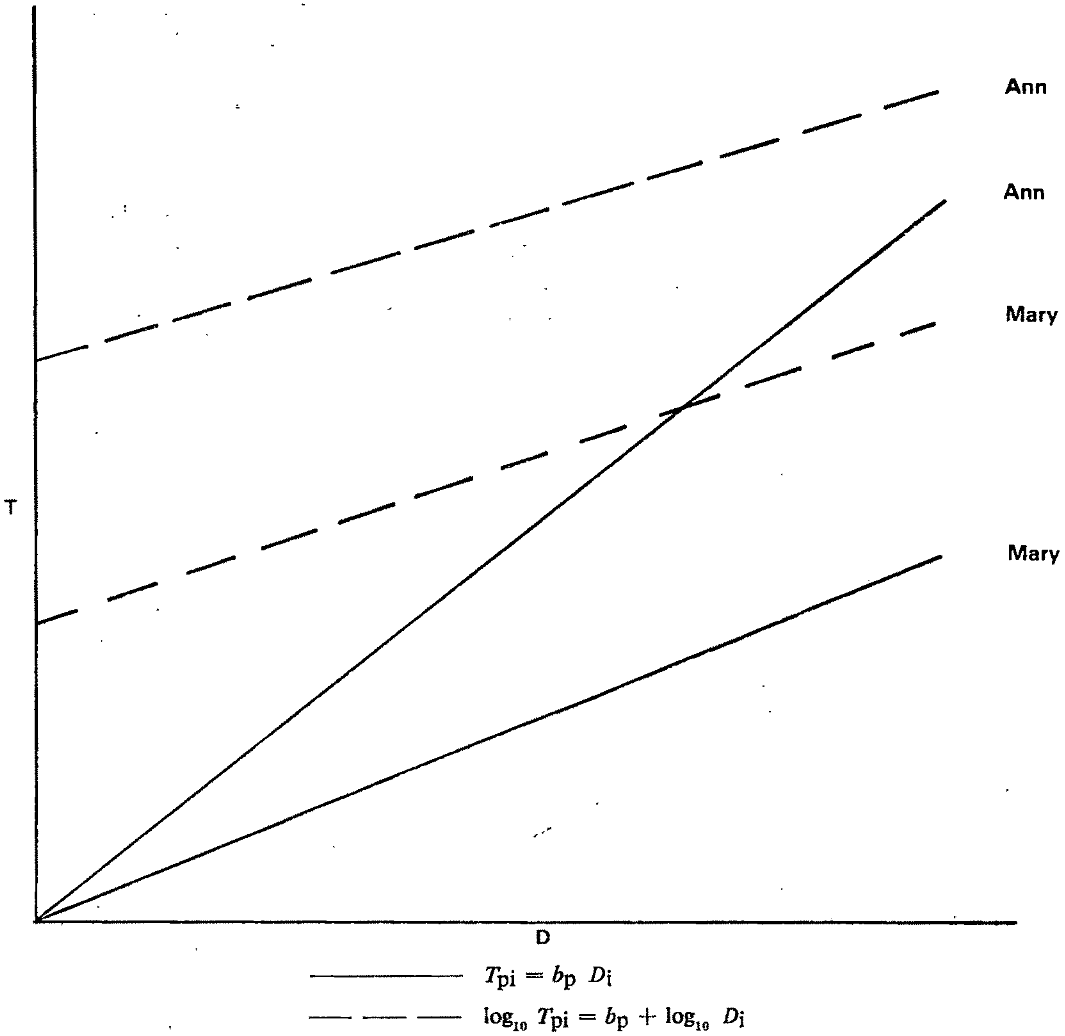
It should be noted that our Model II is directly equivalent to the Furneaux logarithmic model (Furneaux, 1960; Eysenck, 1967). Furneaux logarithmically transformed his item difficulties and individual solution times and found the relationships illustrated by the broken lines in Figure 2. His model takes the form:

$$\log_{10} T_{pi}=b_p+\log_{10} D_i$$

The use of logarithms has, of course, transformed our multiplicative model to an additive model. Thus, in this model, all slopes for individuals are the same, by definition, and individuals are characterised by differences in their intercept on the ordinate.

FIGURE 2

ONE-PARAMETER MODEL



The purpose of the present study is to examine the fit of data to the two models outlined above, and to relate the speed parameters estimated for each individual in the study to a conventional ability criterion and to age.

METHOD

The Test.

A block design test from the British Intelligence Scale (Elliott *et al.*, 1976; Warburton *et al.*, 1970) was employed. The test consisted of 16 items, similar in type to those found in the Wechsler intelligence scales. Children constructed designs presented on cards, using blocks similar, but not identical to Koh's blocks. There was no time limit either for individual items or for the test as a whole.

Sample.

A total of 395 children took the test. Most of the sample were given the test during the try-out phase of the BIS project. They were drawn from all regions of England, Scotland and Wales and were sampled according to criteria designed to make them a reasonably representative sample of the population, although not as much care was taken in sampling as in the standardisation phase currently taking place. The remainder of the sample were children tested for two small-scale research projects at Manchester University.

The children in the sample were aged between 10 and 13 years. They were split into four separate year groups. Additionally, each year group was separately split as evenly as possible into three ability groups on the basis of the total number of items correctly solved on the Block Design test. The 118 children in the 10-year-old group were divided into 48 high-ability children, with raw scores of 14 to 16 on the 16-item test; similarly, there were 48 medium-ability children (raw scores 9 to 13), and 22 low-ability children (raw scores 4 to 8). The 105 11-year-old children comprised 38 of high ability (score 14 to 16), 40 of medium ability (11 to 13) and 27 of low ability (4 to 10). The 67 12-year-olds comprised 23 of high ability (14 to 16), 25 of medium ability (10 to 13) and 19 of low ability (4 to 9). Finally, the 105 13-year-olds comprised 37 of high ability (15 to 16), 38 of medium ability (12 to 14) and 30 of low ability (4 to 11). To ensure reasonably stable estimates of all individual speed parameters, children with fewer than four correct solutions were dropped from the sample.

Procedure.

All children were individually tested by educational psychologists trained to use the test materials. For each item, psychologists noted whether or not the child's solution was correct, together with the time taken to solution.

As there was a wide range of item difficulty, in the usual sense, in the test, an initial analysis was conducted to establish the item facility levels. From the original 16 items in the test, a total of seven items were selected which were passed by 80 per cent or more of the sample. The facility levels (proportion correct) of the seven items were ·90, ·83, ·88, ·86, ·82, ·88 and ·80 respectively. The time difficulties, these being the mean times to correct solution for the total sample, were then calculated for each of the selected items. These, with standard deviations in parentheses, were 15·7 (14·5), 34·6 (25·3), 34·9 (29·5), 43·2 (31·4), 50·6 (31·1), 54·2 (29·9) and 60·3 (29·9) respectively.

For each child, the time to correct solution was regressed upon the time difficulties and the following parameters were estimated:

Model I: a_p and b_p

Model II: b_p

Additionally, for each child, the sum of the squared deviations (over items) of *obtained* times to correct solution from *predicted* times were calculated. These sums were themselves summed over the whole sample and divided by the total number of observations to give the mean squared deviation of the data from the model. This was done separately for models I and II, thus enabling an estimate to be made of any deterioration in fit caused by dropping the a parameter in Model II.

Finally, three two-way analyses of variance were conducted, examining the effects of age (4 levels) and ability (3 levels) upon the two Model I parameters and the single Model II parameter. Because of the unequal cell numbers in the design, the analysis was by the method of unweighted means (Winer, 1962).

RESULTS

Model I.

Summaries of the analysis of variance of the a and b parameters are given in Tables 1 and 2. None of the effects for the a parameter approaches anywhere near a statistically significant level. However, the main effects of both age and ability upon the b parameter are highly significant. Younger and duller children have steeper slopes than older and brighter children. There is no interaction shown between age and ability.

TABLE 1

ANALYSIS OF VARIANCE OF a PARAMETER, MODEL I.

Source	Mean Square	d.f.	F
Age	797·4	3	1·32
Ability.................	277·1	2	·46
Age × Ability	576·6	6	·96
Within groups (error) ..	602·6	383	—

TABLE 2

ANALYSIS OF VARIANCE OF b PARAMETER, MODEL I.

Source	Mean Square	d.f.	F
Age	6·27	3	13·27 P< ·0001
Ability	6·61	2	14·00 P< ·0001
Age × Ability	·27	6	·57
Within groups (error)	·47	383	—

Means of significant effects	Age				Ability		
	10	11	12	13	Low	Medium	High
	1·45	·96	1·03	0·85	1·32	1·03	·86

The a parameter showed marked instability. Values of a for individual children were mainly in the range of ±80, with some being even more extreme, and the parameter showed a high level of within-cell variance, as can be seen in Table 1.

The a and b parameters were highly correlated, the correlations being —0·77, —0·80, —0·80, and —0·88 for the 10, 11, 12 and 13-year-old age groups respectively. Steep slopes were associated with large negative intercepts, and *vice versa.*

Model II.

With the a parameter fixed at zero for all children and thus dropped from the model, only the b parameter remains. The summary of the analysis of variance of this parameter given in Table 3 indicates that the main effects of age and ability are highly significant, with greatly increased F values when compared with those in Table 2. Once again, there is no age × ability interaction effect.

The means for each level of the main effects again indicate that for a given increase in time difficulty, younger and duller children show a relatively large increase in time to correct solution in comparison with older and brighter children.

TABLE 3

Analysis of Variance of *b* Parameter, Model II.

Source	Mean Square	d.f.	*F*
Age	5·16	3	36·92 P< ·0001
Ability	8·13	2	58·15 P< ·0001
Age × Ability	·20	6	1·45 P= ·19
Within Groups (error)	·14	383	—

Means of significant effects	Age				Ability		
	10	11	12	13	Low	Medium	High
	1·32	·96	·99	·74	1·28	·99	·76

Dropping the *a* parameter from Model II appears to have resulted in greater stability of the *b* parameter, this being reflected in two ways. Firstly, the total *b* variance was reduced in Model II compared with Model I (total mean squares being 0·220 and 0·545 respectively). Secondly, the within-cell variance was much reduced, as can be seen from Tables 2 and 3.

The mean squared deviations between observed and expected times for Models I and II were calculated for the whole sample, from a total of 2,539 data points, and were 434·19 and 482·34 respectively. Thus, Model II, by dropping one parameter, has a mean squared deviation only 11 per cent greater than the two-parameter Model I. This lends further support to the view that the *a* parameter can safely be dropped from Model I.

DISCUSSION

The present study provides an encouraging demonstration that a simple theoretically based model for the analysis of speed of problem-solving works better than a more complex empirical model. The single parameter Model II is a special case of Model I, and so could clearly never better its two-parameter rival in goodness of fit to the data. The marginal superiority in fit of the two-parameter model could not justify its retention in the face of its breach of parsimony, decreased parameter stability and less justifiable rules of correspondence of model to theory. The simpler Model II is also firmly in accord with common sense in suggesting that people who are slow in solving problems are not slow by a constant amount, but that as problem difficulty increases, the difference between slow and fast people becomes greater. Simple proportionality appears to be adequate to explain this relation.

The results lend support to previous evidence reviewed earlier and particularly emphasised by Eysenck (1967) that speed is an important component of general intelligence: high test scorers solve problems faster even when no time limits are imposed. Similarly, speed, along with other cognitive abilities, increases as a function of age. The absence of any interaction between age and ability suggests that this increase of speed with age takes place at much the same rate in children of all ability levels. The slight rise in the mean slope of children in the 12-year-old group in comparison with the 11-year-olds may well be due to sampling effects, as there seems to be no theoretical reason for the rise at that age. It should be noted here that the 12-year-old sample was comprised solely of BIS try-out children and was the only age group which did not have additional children pooled from other sources.

It is tempting to think, on the lines of Eysenck (1967), that what may be involved in problem solving activity is some kind of scanning, information processing, or decision-making mechanism, the speed of which will determine the speed of production of the correct solution. Each problem in cognitive tests has a number of facets which require comparison, processing and decision-making activity, each part of which has an associated latency. The sum of these latencies will be the solution time to the problem, more difficult problems requiring a larger number of comparisons or decisions. Hence, for more difficult problems, individuals with slower decision-making latencies would show a greater rate of increase in time to solve them than individuals with relatively fast decision-making latencies.

A similar theory to this has been expounded by Oldfield (Oldfield, 1966; Oldfield and Wingfield, 1964, 1965) and later developed by Lachman (1973, 1974), among others. Oldfield's theory relates to the time taken to name pictures, words which are more frequent in the language producing shorter naming latencies than infrequent words. The theory states that the internal lexical memory store is so arranged that frequent words can be identified through a scanning or binary decision mechanism sooner than infrequent words. Elliott (1976) has shown that while ESN(S) children have longer picture-naming latencies than ESN(M) children for very frequent words—both groups being matched for picture vocabulary knowledge—the less intelligent ESN(S) group showed a greater relative increase in their latency in naming less frequent words. For these words there would theoretically be a larger number of decisions required to identify the name, this having a differential effect upon the two groups of children. The results of that experiment are quite consistent with the present findings.

For the purpose of practical individual measurement of speed, the Block Design test employed in this experiment is hardly ideal. The number of items was small so the error of measurement would be relatively high. The major purpose of the test is to measure individual differences in ability to solve such problems, and the requirement of a speed test in needing a large number of items which everybody can pass is incompatible with the major aim. Consequently, for future measurement of speed in the BIS, 40 items specifically devised for the purpose will be used. These items, consisting of very easy and homogeneous number-checking tasks, are currently being administered as part of the BIS standardisation. It is expected that this test will be analysed in terms of the model outlined in this paper, but with the additional advantage of a reduced error of measurement.

REFERENCES

BAXTER, B. (1941). An experimental analysis of the contributions of speed and level in an intelligence test. *J. educ. Psychol.*, 32, 285-296.

DAVIDSON, W. M., and CARROLL, J. B. (1945). Speed and level in time-limit scores—a factor analysis. *Educ. psychol. Measur.*, 5, 411-427.

ELLIOTT, C. D. (1976). Factors influencing the response latencies of retarded children in naming pictures. (Unpublished manuscript).

ELLIOTT, C. D., PEARSON, L. S., and MURRAY, D. J. (1976). The British Intelligence Scale: final report before standardisation, 1975-76. *Occasional Papers of the Division of Educational and Child Psychology of the British Psychological Society*, 10.

EYSENCK, H. J. (1967). Intelligence assessment: a theoretical and experimental approach. *Br. J. educ. Psychol.*, 37, 81-96.

FURNEAUX, W. D. (1956). *The Nufferno Manuals of Speed and Level Tests.* Published by Institute of Psychiatry, University of London. Distributed by NFER, Slough, Bucks.

FURNEAUX, W. D. (1960). Intellectual abilities and problem solving behaviour. In EYSENCK, H. J. (Ed.), *Handbook of Abnormal Psychology*. London: Pitman.

GUILFORD, J. P. (1956). *Fundamental Statistics in Psychology and Education.* New York: McGraw-Hill.

LACHMAN, R. (1973). Uncertainty effects on time to access the internal lexicon. *J. exp Psychol.*, 99, 199-208.

LACHMAN, R., SHAFFER, J. P., and HENNRIKUS, D. (1974). Language and cognition: effects of stimulus codability, name-word frequency and age of acquisition on lexical reaction time. *J. verb. Learn. verb. Beh.*, 13, 613-625.

MCFARLAND, R. A. (1930). An experimental study of the relationship between speed and mental ability. *J. gen. Psychol.*, 3, 67-97.

OLDFIELD, R. C. (1966). Things, words and the brain. *Q. Jl. exp. Psychol.*, 18, 340-353.

OLDFIELD, R. C., and WINGFIELD, A. (1964). The time it takes to name an object. *Nature*, 202, 1031-1032.

OLDFIELD, R. C., and WINGFIELD, A. (1965). Response latencies in naming objects. *Q. Jl. exp. Psychol.*, 17, 273-281.

PEAK, H., and BORING, E. G. (1926). The factor of speed in intelligence. *J. exp. Psychol.*, 9, 71-94.

RASCH, G. (1960). *Probabilistic Models for Some Intelligence and Attainment Tests.* Copenhagen: Danish Institute for Educational Research.

SLATER, P. (1938). Speed of work in intelligence tests. *Br. J. Psychol.*, 29, 55-68.

SLATER, P. (1944). Scores of different types of neurotics on tests of intelligence. *Br. J. Psychol.*, 35, 40-42.

STUDMAN, G. (1935). The measurement of speed and flow of mental activity. *J. ment. Sci.*, 81, 107-137.

SUTHERLAND, J. D. (1934). The speed factor in intelligent reactions. *Br. J. Psychol.*, 24, 276-294.

TATE, M. W. (1950). Notes on the measurement of mental speed. *J. educ. Psychol.*, 41, 219-236.

THORNDIKE, E. L. (1926). *The Measurement of Intelligence.* New York: Teachers' College, Columbia University.

THURSTONE, L. L. (1938). *Primary Mental Abilities.* Chicago: University of Chicago Press.

VAN DER VEN, A. H. G. S. (1971). Time-limit tests: a critical evaluation. *Ned. Tijdschr. Psychol.*, 26, 580-591.

VAN DER VEN, A. H. G. S. (1972): Speed and precision in the General Aptitude Test Battery. *Ned. Tijdschr. Psychol.*, 27, 530-537.

VAN DER VEN, A. H. G. S. (1973a). *A Theory of Time-limit Tests.* Report 73 MA10, Department of Psychology, Nijmegen University.

VAN DER VEN, A. H. G. S. (1973b). A true-score theory for time-limit tests. *Ned. Tijdschr. Psychol.*, 28, 267-280.

VAN DER VEN, A. H. G. S. (1974). *The Correlation between Speed and Precision in Time-limit Tests.* Report 74 MA02, Department of Psychology, Nijmegen University.

WARBURTON, F. W., FITZPATRICK, T. F., WARD, J., and RITCHIE, M. (1970). Some problems in the construction of individual intelligence tests. In MITTLER, P. (Ed.), *The Psychological Assessment of Mental and Physical Handicap.* London: Methuen.

WHITE, P. O. (1973). Individual differences in speed, accuracy and persistence: a mathematical model. In EYSENCK, H. J. (Ed.), *The Measurement of Intelligence.* Lancaster: Medical and Technical Publ. Co.

WINER, B. J. (1962). *Statistical Principles in Experimental Design.* New York: McGraw-Hill.

(*Manuscript received* 21*st April*, 1976)

Br. J. educ. Psychol., 47, 60-65, 1977

LIP-READING BY DEAF AND HEARING CHILDREN

By R. CONRAD

(*Medical Research Council Applied Psychology Unit, Cambridge*)*

SUMMARY. A group of profoundly deaf 15-year-old subjects with no other handicap and of average non-verbal intelligence were given a lip-reading test. The same test was given to comparable hearing subjects ' deafened ' by white noise masking. The difference between the groups was not significant. Performance of the deaf subjects (and the hearing) was, however, significantly better when they read the test items from print. The implication is that the relatively poor lip-reading ability of the deaf subjects was not because of linguistic impairment.

INTRODUCTION

THE education of profoundly deaf children in Britain, more so than in many other countries, is firmly committed to an ' oral ' philosophy. A principal objective is to provide the skills which will permit a deaf child to communicate with hearing people in a natural manner. With this objective there is no alternative to ordinary speech on the part of the deaf child. But he must also be able to understand speech. Hood and Poole (1971) have reported that when the degree of hearing loss is as slight as 31dB discrimination of speech sounds begins to be impaired. At this level of hearing loss understanding of speech is very nearly normal and most such children will be educable in ordinary schools, especially when provided with appropriate hearing aids.

As hearing becomes more impaired, even with a hearing aid in perfect working order, a point is reached where effectively none of the speech sounds which might be present in normal conversation is of usable intelligibility. At this point, almost all of the child's understanding of speech depends on visible features; the child must lip-read. Clearly no precise value can be given to represent a level of hearing loss at which no words are identifiable by ear in ordinary circumstances. Children differ, listening conditions differ, speakers differ, quality, appropriateness and servicing of hearing aids all differ—to mention a few of the relevant factors. But it is reasonable to assume that when hearing loss reaches a value of around 95dB (ISO), a child must lip-read if he is to understand speech. Hine (1973) has reported that about 50 per cent of children in England and Wales receiving special education because of deafness are in this category.

With this degree of deafness, a child whose lip-reading is poor must have recourse either to esoteric modes of communication or remain linguistically mute. In an educational system relying upon oral communication, a child with poor lip-reading, even if he has acquired manual signing extra-curricularly, will have few people with whom he can communicate; principally others who are deaf. Since the vast majority of deaf children have normally hearing parents and siblings (Gallaudet College, 1973) there will be a great many very deaf children, possibly the majority, who, without adequate lip-reading, would have extremely limited communication even within their own family.

---

* Now at the Department of Experimental Psychology, University of Oxford.

The social logic of oral education for the deaf is, therefore, impeccable. Profoundly deaf children must be taught to lip-read whether or not other communication modes are also taught. However, the social imperative is independent of the question of effectiveness and it is pertinent to ask how effective lip-reading is by the time a child leaves school.

Very few published studies have evaluated the lip-reading ability of profoundly deaf children. The Department of Education and Science (1964, 1972) has presented some discouraging subjective assessments based on conversations, and Montgomery (1968) some inter-relationships of lip-reading with other abilities. A few studies have examined the relative effectiveness of lip-read speech when compared with other communication modes such as print-reading, manual-sign-reading and speech-sign combinations (Gates, 1971; Klopping, 1971; White and Stevenson, 1975). In general, lip-reading is shown to be a disadvantageous means of understanding verbal information. But, apart from Department of Education and Science (1964, 1972), none of these studies provides information in an absolute sense about the effectiveness of lip-reading for understanding speech. Lip-reading may be better than some other modes and one child may be better than another (Craig, 1964), but as a communication skill for use with hearing people, little is known about performance levels.

Unfortunately, the routine application of a lip-reading test does not meet the problem. A child who achieves a perfect score might perform badly on a different test which used different vocabulary and syntax. At best it will indicate relative performance amongst deaf children. The problem is quite different from that of assessing prose comprehension for instance, where a reading-age, because of the pragmatics of wide-ranging experience, has a value which is meaningful with respect to the objectives of the skill. So at present it is simply not known what a deaf lip-reader is capable of understanding except against other communication modes and other lip-readers. Nor is it easy to see how such knowledge could become available objectively. Yet the crucial problem of performance remains.

In an attempt to evade this *impasse* and still try to obtain information about absolute levels of usable lip-reading skill, the present study compared the lip-reading ability of profoundly deaf children with that of ' deafened ' hearing children. This would be a comparison where the deaf group had had extensive training and practice in lip-reading, whilst the hearing group would effectively have none at all. Thus, on what seems to be a reasonable assumption that formal training and wide experience would be reflected in performance, one would have to predict that hearing children would be very substantially worse lip-readers.

METHOD

Sample. 67 deaf children were drawn variously from amongst nearly all of the special schools for the deaf in England and represent close to an 80 per cent sample of a population with the following characteristics: all subjects were aged 15—16½ years at the time of testing; were known to be deaf prior to the age of 3 years ; had a hearing loss in the better ear of greater than 94dB averaged across the frequencies 250—4,000 Hz; had no other reported handicap; ' spoke ' English at home; had received special education because of deafness, including lip-reading, for at least 10 years. The final group of 67 subjects was reached after excluding some mentally retarded children who otherwise met the above criteria. Montgomery (1968) has shown a significant correlation between lip-reading and score on the Progressive Matrices. All subjects, therefore, were given the Standard Progressive Matrices (Raven, 1960) and the mean of the normally distributed scores was 44, which represents the average non-verbal intelligence of the hearing population of the same age. The deaf group then, apart from being deaf, were otherwise normal. Approximately half were boys.

Two groups of hearing subjects were used, one of 75 children, the other of 43. All were drawn randomly from children aged 15–16 years in a medium-sized comprehensive school in a semi-rural area. It was assumed that the average intelligence of both groups was normal. None of these children required a hearing aid.

Test. An amended version of the Donaldson Lip-Reading Test (Montgomery, 1966) was used. This is a face-to-face live-speech test, widely used in Scotland, and the only such available test known to the author and linguistically suitable for British children. The test is of the following form. The speaker makes a short statement and the subject points to one of a number of photographs (6–9), only one of which is correct. Statements are of the kind: " a cat and a dog "; " the woman with dark hair wears glasses." The original test contains 40 such statements.

Because the test was intended for use with children of all ages, the present author had found that it did not sufficiently discriminate amongst older children because of ceiling effects. A further 10 statements were added which successfully increased the overall difficulty of the test. It was also found that a number of test items gave inconsistent results, and when these were eliminated from the score (though not from the test) the maximum score became 38. The test takes about 10 minutes.

Procedure. Deaf children were tested individually, in good illumination with speakers directly facing at a distance of 2–3 feet, and with the pictorial answer booklet between them. This is a situation with which deaf children are quite familiar. For the present purpose, the statements were made clearly but in a normal conversational tone. No statements may be repeated and the rules are clarified during preliminary practice.

Immediately following lip-reading, subjects again did the test, but this time the same statements were typed on individual cards. The subject read each statement and again pointed to the answer. This was done because in order to lip-read a statement, a number of processes are involved, some of which are visual and some linguistic. A failure in visual analysis of lip and tongue positions will result in a wrong answer. But a child who makes a correct analysis but fails to understand the vocabulary or syntax used may also answer incorrectly. It is then debatable whether failure to respond correctly represents a true lip-reading failure. The print version was an attempt to examine this feature of performance and possibly overcome the confounding effects of linguistic knowledge on lip-reading ability. Subjects were not told whether their answers were correct.

The group of 75 hearing children were given the test in lip-read form twice in a single session. The first run was regarded as practice and discarded from the results. No training was given and no other practice. The hearing subjects wore headphones through which was passed white noise at an intensity sufficient to mask the sound of conversational speech at 2 feet. This is achieved with relatively low intensities and is nowhere near enough to mask the sound of one's own voice. The technical difficulties of measuring signal-noise ratios at the ear were simply ignored, but regular checks were made by unexpectedly speaking to the child with lips concealed. It can reasonably be taken for granted that the hearing subjects were completely 'deafened,' receiving no auditory speech information at all from the speaker. Again, knowledge of results was withheld.

Because it was felt that it would strain motivation too far to give three administrations of the same test to the same subjects, a different group of 43 hearing children drawn randomly from the same population did the print version just once. It was assumed once again that this group averaged normal intelligence and was equivalent in linguistic knowledge to the first hearing group. These children were tested without noise masking, but in otherwise identical conditions.

Speakers. Three female speakers were used at different times. All had taught in schools for the deaf for several years. Preliminary trials were conducted to examine the ease with which deaf children could lip-read all three speakers. The trials used deaf subjects not involved in the present study, but did use the same test. Speakers, in pairs, both faced the subject and spoke the test statements alternately with appropriate balancing. No differences between speakers could be detected in the results. In the main study, therefore, the speakers were used as convenient.

RESULTS

Table 1 gives the mean score for the three groups reading lips and reading print.

TABLE 1

MEAN SCORES ON THE AMENDED DONALDSON LIP-READING TEST AND PRINT VERSION.

	Condition	
	Lip-reading	Print reading
Deaf	Mean (SD) 26·95 (8·24)	Mean (SD) 33·52 (5·01)
	Same group N=67	
Hearing	25·30 (5·76) N=75	36·70 (1·39) N=43

(Maximum score=38)

On lip-reading, the small difference between deaf and hearing children is not significant (t=0·90). The print version does show a difference, but this has little relevance. It merely confirms that profoundly deaf children are linguistically retarded when compared with hearing children. For the deaf group, the difference between lip-reading and print-reading is highly significant (t=9·76; P<0·001) and this would obviously be the case for hearing children.

DISCUSSION

These results pose the following problem. A randomly selected group of hearing children with no more than 10 minutes of practice, with no instruction at all, deprived of all auditory speech information and in a totally unfamiliar situation, nevertheless, lip-read as well as a group of 'normal' deaf children who have been exposed to lengthy formal instruction, and with many years of experience. Doubtless, were the comparison made using deaf children with less hearing impairment, their performance would have been better. But it would be pointless to compare deaf subjects with some residual hearing for speech sounds with completely deafened hearing subjects. The problems of providing hearing subjects with matched ' residual hearing ' are beyond current technology. So because it is virtually certain that the hearing subjects heard nothing whilst lip-reading, whilst some of the deaf children may have had a little useful residual hearing, the comparison may be biased in favour of the deaf. But it is also possible that there may have been a different bias in favour of the hearing group. The practice enjoyed by the deaf had extended over many years—but it was generalised over all lip-reading situations. The hearing group had specific practice on the test used—though of very short duration. On the other hand the hearing subjects were in a more unusual situation, whilst the deaf were not. The significance of these effects cannot be quantified and no obvious alternative procedure was available.

The questions at issue are: how effective is unaided lip-reading as a communication mode? How well can deaf children lip-read? It was noted at the beginning of this article that absolute answers to these questions cannot be readily to hand. But for both groups it is clear that far less is communicated through lip-reading than through reading print. Lip-reading is not as effective as reading print and there are well understood reasons to account for this, the principal being that far too many phonemes are not visible on the lips (Erber, 1974a). But this means that by lip-reading, the deaf child cannot exploit the knowledge of language that he does have and which is, in any case, relatively impoverished.

But in a sense the more interesting question is the second, concerning performance. Here it seems that performance is no better than that of inexperienced hearing children. It is hard to evade the conclusion either that almost all a child can learn about lip-reading can be self-taught in 10 minutes, or that educators do not have the necessary theoretical understanding of the processes involved to be able, usefully, to intervene. It is a fact that whilst there are a number of studies concerned with the physical conditions for aiding lip-reading, such as illumination and distance factors (Erber, 1971, 1974b), it is difficult to cite a single experimental study aimed at elucidating the cognitive factors which must underly the process. In this context it is particularly important to distinguish the cases of the adventitiously deafened adult who is already familiar with spoken language, and the child born deaf. The former, in learning to lip-read, translates from a second language to a first—which he already knows. For the deaf subjects of this study, lip-reading is generally regarded as the first language equivalent to heard speech for the hearing child. The two cases are, therefore, not only not comparable, but in many aspects misleading.

It is because the deaf child has no mother-tongue into which lip-read patterns can naturally be translated that it was also necessary to determine the linguistic understanding of the deaf group for the test items. The print version results confirm the expectation that the hearing subjects would have the meanings of the test items available, requiring only translation from lip-reading, to achieve comprehension. It was far from certain that this would be the case for the deaf subjects. Their lip-reading score might have faithfully represented the limit of their linguistic knowledge. It would then have been possible to say that because deaf subjects could lip-read whatever they knew linguistically—not the case with hearing subjects—lip-reading was an effective means of communication, limited only by linguistic knowledge. There would be no evidence to show that deaf children were not good lip-readers. In this case, educational effort would properly be directed towards improving knowledge of language.

In the event the results show that lip-reading is frequently unable to extract meaning which is within the subject's competence and available through reading. The comparison of hearing subjects tested on clear speech and reading was dispensed with because the outcome, in view of the brevity and simplicity of the test items, seemed too obvious.

The development of speech in man depends on the ability to hear. Words do not sound alike unless their distinct meanings can be derived from linguistic context. There is no reason why the development of speech should make it a suitable communication mode in the absence of correlated sound. Promoting lip-reading for this purpose is, in a sense, a shot in the dark. The shot depends substantially on what of speech is visible and about which little can be done, and upon what can be taught with respect to making use of what is visible. The evidence of the present study suggests that education of profoundly deaf children is at present theoretically ill-equipped to teach English (at least) as a first language by means of lip-reading. Whether this conclusion is an argument for abandoning the attempt on the grounds of the unsuitability of the spoken form of the language—implying the need for a replacement language—seems to remain a wide open question.

ACKNOWLEDGMENTS—I am indebted to Dr. G. W. G. Montgomery for providing the Donaldson Lip-Reading Test, and to Dr. J. G. Kyle for making the necessary amendments. Testing was shared by Mrs. B. C. Weiskrantz, Miss M. G. McKenzie and Miss A. J. M. Morris.

REFERENCES

CRAIG, W. N. (1964). Effects of pre-school training on the development of reading and lip-reading skills of deaf children. *Amer. Ann. Deaf.*, 109, 280-296.

DEPARTMENT OF EDUCATION AND SCIENCE (1964). *The Health of the School Child*, 1962 and 1963. London: HMSO.

DEPARTMENT OF EDUCATION AND SCIENCE (1972). *The Health of the School Child*, 1969-70. London: HMSO.

ERBER, N. P. (1971). Effects of distance on the visual reception of speech. *J. Speech Hear. Res.*, 14, 848-857.

ERBER, N. P. (1974a). Visual perception of speech by deaf children : recent developments and continuing needs. *J. Speech Hear. Dis.*, 39, 178-185.

ERBER, N. P. (1974b). Effects of angle, distance and illumination on visual reception of speech by profoundly deaf children. *J. Speech Hear. Res.*, 17, 99-112.

GALLAUDET COLLEGE (1973). *Characteristics of Hearing Impaired Students by Hearing Status.* Washington, D.C.: Office of Demographic Studies, Gallaudet College.

GATES, R. R. (1971). The reception of verbal information by deaf students through a television medium—a comparison of speech-reading, manual communication, and reading. *Proc. Conv. Amer. Instr. Deaf*, 513-522.

HINE, W. D. (1973). How deaf are deaf children? *Br. J. Audiol.*, 7, 41-44.

HOOD, J. D., and POOLE, J. P. (1971). Speech audiometry in conductive and sensorineural hearing loss. *Sound*, 5, 30-38.

KLOPPING, H. W. E. (1971). Language Understanding of Deaf Students under Three Auditory-Visual Stimulus Conditions. Unpublished Ed.D. Diss., University of Arizona.

MONTGOMERY, G. W. G. (1966). The relationship of oral skills to manual communication in profoundly deaf adolescents. *Amer. Ann. Deaf*, 111, 557-565.

MONTGOMERY, G. W. G. (1968). A factorial study of communication and ability in deaf school leavers. *Br. J. educ. Psychol.*, 38, 27-37.

RAVEN, J. C. (1960). *Guide to the Standard Progressive Matrices.* London: Lewis.

WHITE, A. J., and STEVENSON, V. M. (1975). The effects of total communication, manual communication, oral communication and reading on the learning of factual information in residential school deaf children. *Amer. Ann. Deaf.*, 120, 48-57.

(Manuscript received 5th March, 1976)

RESEARCH NOTES

Br. J. educ. Psychol., 47, 66-70, 1977

EFFECTS ON CHILDREN'S DIVERGENT THINKING ABILITIES OF A PERIOD OF DIRECT TEACHING FOR DIVERGENT PRODUCTION

By B. S. FRANKLIN
(Gloucestershire College of Education)
AND P. N. RICHARDS
(Bath University, School of Education)

SUMMARY. Teaching procedures designed to improve children's divergent thinking abilities were administered to two classes of 9- and 10-year-old children, one formal and one informal, for one morning per week over a period of six weeks. Parallel classes covered the same material but without the emphasis on divergent production. The results suggest that appropriate teaching procedures can play a significant part in enhancing children's divergent thinking abilities, particularly verbal abilities. Similar improvements were attained in both types of experimental class.

INTRODUCTION

Although divergent thinking abilities have been investigated in this country in a variety of studies, there has been little investigation into the effect on such abilities of direct teaching for divergent production. Explorations have been made into the effects on children's divergent thinking of differing teaching approaches (Haddon and Lytton, 1968, 1971), school climate (Lytton and Cotton, 1969), type of mathematics teaching (Richards and Bolton, 1971), and curriculum structure (Ogilvie, 1974) but each of these was concerned with the cumulative effects of some years of schooling rather than specific training.

Apart from the work of de Bono (1973) in creating materials for ' teaching thinking directly,' most of the attempts to improve children's creative thinking abilities by using specific teaching programmes have been made in the USA, based on the conviction that " like most behaviour, creative activity probably represents to some extent many learned skills " (Guilford, 1952) Summarising 27 such attempts in which the major criteria were tests of divergent thinking, Torrance (1972) reports that significant improvement was shown on over 90 per cent of the criteria used. Treffinger *et al.*, (1974) investigating the effectiveness over periods of four and eight weeks of two programmes designed to improve children's creative thinking and problem solving ability, found a similar improvement and concluded that " both programmes have been shown to effect significant enhancement of fifth-grade children's divergent thinking abilities (particularly verbal abilities). "Without entering into a discussion of the validity of using divergent thinking tests to evaluate programmes designed to teach children to think creatively, the improvements in divergent thinking are consistent with Guilford's (1967) view that appropriate teaching methods and techniques will promote the development of the different intellectual abilities in his structure of intellect model, including those of divergent thinking. Avoiding the creativity label the teaching in this investigation is aimed directly at improving children's divergent thinking ability.

METHOD

Sample and Design.

119 children of 9 and 10 years of age took part in the whole study. They comprised four third-year classes, one in each of four unstreamed junior schools with similar catchment areas in the suburbs of a small provincial city. After investigating the schools available, it was found possible to form two pairs of classes, each pair similar in age, social class and classroom environment, and it was decided to replicate the experiment within each pair. IQ was measured on the NFER Verbal Test BC and

social class according to the Registrar-General's classification. An attempt was made initially to classify the local schools as formal or informal as Gardner (1966) and Haddon and Lytton (1968) have done, but without the benefit of some recent findings about such classifications (Bennett and Jordan, 1975). It proved impossible, however, as the head teachers concerned favoured aspects of both styles and the schools did not wholly adhere to either category. Within the schools, nevertheless, head teachers did identify certain teachers as formal or informal, and further information was then collected about their classrooms, including the type of pupil-teacher relationship, the arrangement of furniture and material on display. This finally enabled classes with a similar atmosphere to be paired and allows the effects of the teaching for divergent production to be assessed within different classroom environments. Details of the sample are given in Table 1.

TABLE 1

DESCRIPTION OF SAMPLE.

	Age		IQ		Social Class (Percentage distribution)			Classroom Environment
	Mean	SD	Mean	SD	I/II	III	IV/V	
Class A .. (N=29)	9y 10m	3·3m	102·5	12·1	20	58	22	Formal
Class B .... (N=32)	9y 10m	4·0m	106·6	9·3	21	56	23	Formal
Class C .... (N=29)	9y 10m	3·8m	96·7	10·8	20	60	20	Informal
Class D .... (N=29)	9y 10m	3·7m	98·1	10·6	20	55	25	Informal

(There are no significant differences in Age, IQ or Social Class within either pair.)

Divergent thinking tests.

Seven tests were used, five verbal and two figural, each separately timed. In all cases the items in the tests were changed for the post-test. The tests were: Alternate Uses (4 items, 15 mins.), Similarities (4 items, 15 mins.), Line Meanings (4 items, 10 mins.), Instances (1 item, 10 mins.) all used by Wallach and Kogan (1965); Drawing Completion (10 mins.) (Torrance, 1966) using circles in the pre-test and parallel lines in the post-test; Statements (1 item, 10 mins.) and Questions (1 item, 10 mins.). The two latter tests were adapted from Guilford and Torrance and asked subjects to list things they could say about, and ask about, a given object.

Each test was scored for Fluency and Originality. The originality score follows Torrance (1966), though it is in fact a score for statistical infrequency rather than uniqueness and might more appropriately be termed unusualness as suggested by Vernon (1971). In presenting the test the children were asked to think of as many things as possible and also unusual things. They were asked to work in silence though the game-like nature of the tests was emphasised. They were provided with drawing paper and crayons to occupy them if they finished early. For each class the testing took place in two sessions of 40 and 30 minutes respectively.

Procedure.

Each class was taught for one morning per week, for 10 weeks, by one of the writers. The experimental classes (A and C) carried out the divergent production exercises over a period of six weeks, the two weeks preceding and following this being used for the testing procedures. The control classes (B and D) followed a similar pattern, but the six-week period was devoted to artistic work and more formal lessons covering the same material as the experimental classes but without the emphaiss on thinking and finding out for themselves that characterised the experimental procedures. In order, however, that they should find the work as enjoyable and novel as the experimental classes, the art lessons concentrated on providing a variety of new and stimulating experiences.

The divergent production exercises.

The exercises were aimed at developing the children's divergent production ability—to generate information where the emphasis is upon variety and quantity of output related to a given source. As Guilford (1959) notes such ability involves " a type of thinking in which considerable searching about is done," and the work aimed to develop both the skills and the motivation for such thinking. The activities emphasised the skills of questioning and finding out and encouraged the children to view their learning as a personal activity in which they should think for themselves. In response to the presentation of a stimulus object, information was generated by the children from previous knowledge, class discussion, experiments and individual research and reference. A categorisation scheme was presented to help structure the work.

The subject of the first morning was ' bricks,' a brick having been the object used previously in the Statements and Questions tests. Two charts were displayed on which the children's response to these tests were shown in six categories, descriptions, components, classifications, origins, prognoses and applications. These correspond to the questions, how can we describe it, what is it made of, what type of thing is it and what other kinds are there, where does it come from, what can happen to it, and what can be done with it? In the subsequent discussions, additional responses were added and also as many answers to the questions as the children could suggest. A number of questions remained unanswered and the children were asked to choose from these and others suggested by the experimenter, and carry out an investigation and communicate the information gained. Some of the activities were experimental, e.g., how heavy is a brick, how many are needed to build a wall 6 feet long by 4 feet high? Others involved searching for information, e.g., when were bricks first made, how are bricks made today? The findings were presented in drawings, writings, models, maps, graphs, etc. Questions which remained unanswered were left displayed and during subsequent weeks answers were added whenever information was brought in by the children. In the last part of the morning a new object—an apple, was presented to the class and lists made of all the things the children could say about it and all the questions they could ask about it. Finally, following up the statement that the apple was round, a list of other round things was compiled. A similar teaching procedure was followed in subsequent weeks, information being generated about apples, coins, things made in moulds, things that burn and types of round objects.

RESULTS

The mean gain of each class and the difference between the performances of the experimental and control classes is given in Table 2 for each of the divergent thinking measures. Significance is assessed by the two-tailed t-test. Although a one-tailed test would have been appropriate where gains were expected, a two-tailed test has been applied throughout for consistency. The results are still very significant. Both experimental classes improved their divergent performance on every variable, the formal class showing significant gains on nine of out 14 measures and the informal class on seven. The control classes showed both gains and losses, each class improving significantly on three measures and deteriorating on two. In the formal pair the experimental class had a higher score than the control on every variable, significantly so in nine cases. In the informal pair the experimental class had a higher score than the control class on each verbal test, significantly so in seven cases, but improved less than the control class on the four figural tests.

TABLE 2

MEAN GAIN SCORES OF EACH CLASS ON EACH OF THE DIVERGENT THINKING MEASURES AND A COMPARISON OF EACH PAIR OF CLASSES.

Group		Verbal Tests										Figural Tests			
		Alternate Uses		Similarities		Statements		Instances		Questions		Line Meaning		Drawing Completion	
Formal Classes		flu.	orig.	flu.	orig.	flu.	orig.	flu.	orig.	flu.	orig.	flu.	orig.	flu.	orig.
CLASS A (Experimental)	mean gain	3·65	4·13	4·37	4·03	6·75	3·44	5·34	5·62	2·10	0·58	1·58	2·55	4·13	1·72
	t	3·38	1·92	5·46	3·39	7·42	1·87	3·66	3·10	2·88	0·68	2·08	1·73	4·80	1·09
	P<	·01	NS	·001	·01	·001	NS	·001	·01	·01	NS	·05	NS	·001	NS
CLASS B (Control)	mean gain	0·68	–0·62	0·75	–3·59	2·37	0·37	–0·75	–1·03	–0·06	–1·96	1·46	0·93	1·96	–1·53
	t	0·96	0·46	1·03	3·82	4·16	0·60	0·82	0·95	0·12	2·45	2·52	0·89	3·38	1·20
	P<	NS	NS	NS	·001	·001	NS	NS	NS	NS	·05	·05	NS	·01	NS
Difference between mean gains	diff. (A–B)	2·97	4·75	3·62	7·62	4·38	3·07	6·09	6·65	2·16	2·54	0·12	1·62	2·17	3·25
	t	2·37	1·90	3·48	5·18	4·21	1·59	3·62	3·21	2·43	2·24	0·13	0·91	2·17	1·64
	P<	·05	NS	·01	·001	·001	NS	·01	·01	·05	·05	NS	NS	·05	NS
Informal Classes															
CLASS C (Experimental)	mean gain	4·37	3·79	6·55	2·48	6·20	2·27	7·44	5·10	3·93	1·72	0·44	1·06	2·79	0·10
	t	2·91	1·71	6·62	1·62	10·33	3·15	4·96	2·91	4·97	1·61	0·54	0·99	3·93	0·10
	P<	·01	NS	·001	NS	·001	·01	·001	·01	·001	NS	NS	NS	·001	NS
CLASS D (Control)	mean gain	–4·24	–3·41	1·93	–0·82	2·24	0·62	1·00	–1·41	1·20	–0·62	1·27	1·65	3·44	0·24
	t	6·73	3·22	2·88	0·88	4·77	1·69	1·30	1·76	1·69	0·76	1·98	1·23	6·04	0·23
	P<	·001	·01	·01	NS	·001	NS	NS	NS	NS	NS	NS	NS	·001	NS
Difference between mean gains	diff. (C–D)	8·61	7·20	4·62	3·30	3·96	1·65	6·44	6·51	2·73	2·34	–0·82	–0·59	–0·65	–0·14
	t	5·31	2·95	3·91	1·85	5·21	1·79	3·85	3·40	2·60	1·75	0·82	0·34	0·72	0·09
	P<	·001	·01	·001	NS	·001	NS	·001	·01	·05	NS	NS	NS	NS	NS

CONCLUSIONS

The success of the divergent production exercises in improving the children's divergent thinking indicates that the latter can be facilitated by teaching procedures that help develop relevant skills and the motivation to use them. This gives support to Guilford's suggestion (1967) that appropriate teaching methods and techniques will promote the development of divergent thinking, and to the view that the improvements are most predictable when deliberate teaching is involved (Torrance, 1972).

Like that of Treffinger *et al.* (1974), the improvement in this study is mainly on the verbal measures of divergent thinking, though this is consistent with the form of the procedures adopted for fostering divergent production. The improvement of both experimental classes irrespective of their formal or informal nature, gives weight to the generality of the results.

REFERENCES

BENNETT, S. N., and JORDAN, J. (1975). A typology of teaching styles in primary schools. *Br. J. educ. Psychol.*, 45, 20-28.

DE BONO, E. (1973). *CORT Thinking: Teachers' Notes.* Blandford, Dorset: Direct Educational Services, Ltd.

GARDNER, D. E. M. (1966). *Experiment and Tradition in Primary Schools.* London: Methuen.

GUILFORD, J. P. (1952). Some recent findings on thinking abilities and their implications. *Inform. Bull.*, 3, 48-61.

GUILFORD, J. P. (1959). Traits of Creativity. In ANDERSON, H. H. (Ed.), *Creativity and its Cultivation.* New York: Harper.

GUILFORD, J. P. (1967). *The Nature of Human Intelligence.* New York: McGraw Hill.

HADDON, F. A., and LYTTON, H. (1968). Teaching approach and the development of divergent thinking abilities in primary schools. *Br. J. educ. Psychol.*, 38, 171-180.

HADDON, F. A., and LYTTON, H. (1971). Primary education and divergent thinking abilities—four years on. *Br. J. educ. Psychol.*, 41, 136-147.

LYTTON, H. and COTTON, A. C. (1969). Divergent thinking abilities in secondary schools. *Br. J. educ. Psychol.*, 39, 188-90.

OGILVIE, E. (1974). Creativity and curriculum structure. *Ed. Res.*, 16, 126-132.

RICHARDS, P. N., and BOLTON, N. (1971). Type of mathematics teaching, mathematical ability and divergent thinking in junior school children. *Br. J. educ. Psychol.*, 41, 32-37.

TREFFINGER, D. J., SPEEDIE, S. M., and BRUNNER, W. D. (1974). Improving children's creative problem solving ability: The Purdue Creativity Project. *J. Creat. Behav.*, 8, 20-30.

TORRANCE, E. P. (1966). *Torrance Tests of Creative Thinking: Technical Norms Manual.* Princeton: Personnel Press.

TORRANCE, E. P. (1972). Can we teach children to think creatively? *J. Creat. Behav.* 6,, 114-143.

VERNON, P. E. (1971). Effects of administration and scoring on divergent thinking tests. *Br. J. educ. Psychol.*, 41, 245-257.

WALLACH, M. A., and KOGAN, N. (1965). *Modes of Thinking in Young Children.* New York: Holt, Rinehart and Winston.

(*Manuscript received* 12*th July*, 1976)

Br. J. educ. Psychol., 47, 71-75, 1977

READING ATTAINMENT AND FAMILY SIZE: AN ANOMALY

BY K. RICHARDSON
(*National Children's Bureau, London*)

SUMMARY. An analysis of the written compositions of a national sample of 520 British school children was carried out using two measures: (i) writing productivity as measured by the composition length in words; (ii) the mean minimal terminable unit (T-unit) length in words, a measure of syntactic maturity. Although this sample exhibited the typical association between family size and reading attainment, there were no family size differences with respect to writing productivity or syntactic maturity. The results were discussed with regard to the possibility of a syntactic differentiation among children from families of larger size such that the syntactic content of standardised reading comprehension tests would prejudice their test performance.

INTRODUCTION

The association between family size and school attainment is a long-established finding of developmental research. In Britain it has drawn further support from two well-known longitudinal studies, the National Survey of Health and Development (e.g., Douglas, 1964) and the National Child Development Study (Davie *et al.*, 1972; Fogelman, 1975; Fogelman and Goldstein, 1976). Two explanations for the relationship have commonly been adopted (see below). The present communication is restricted to a consideration of reading attainment. It illustrates the relationship between family size and reading attainment and draws attention to two sets of collateral data which, in juxtaposition with the former, may subserve an additional or alternative form of explanation.

METHOD

The subjects were 520 11-year-old boys and girls, comprising a randomly selected sub-sample from the National Child Development Study. The latter is a longitudinal study of the approximately 16,000 children in England, Scotland and Wales born in one week of March, 1958 (for details, see Davie *et al.*, 1972). The present data were collected during the second follow-up study when the children were 11 years old (see Davie, 1973). Reading attainment was tested using a sentence-completion reading comprehension test which was constructed as a parallel version of the Watts-Vernon test by the National Foundation for Educational Research. Examples of items are:

(i) " When the speaker asked if there were any questions he was (involved, immured, inundated, implied, instructed) with queries."

(ii) "As we both come from the same town, my wife and I have a great many (typical, mutual, friable, arable, viable) friends."

The test was administered by teachers in schools under strict instructions regarding circumstances and timing. After completion of this test the children were asked to write a composition on what they imagined their life to be like at the age of 25 years. Thirty minutes were allowed for this task.

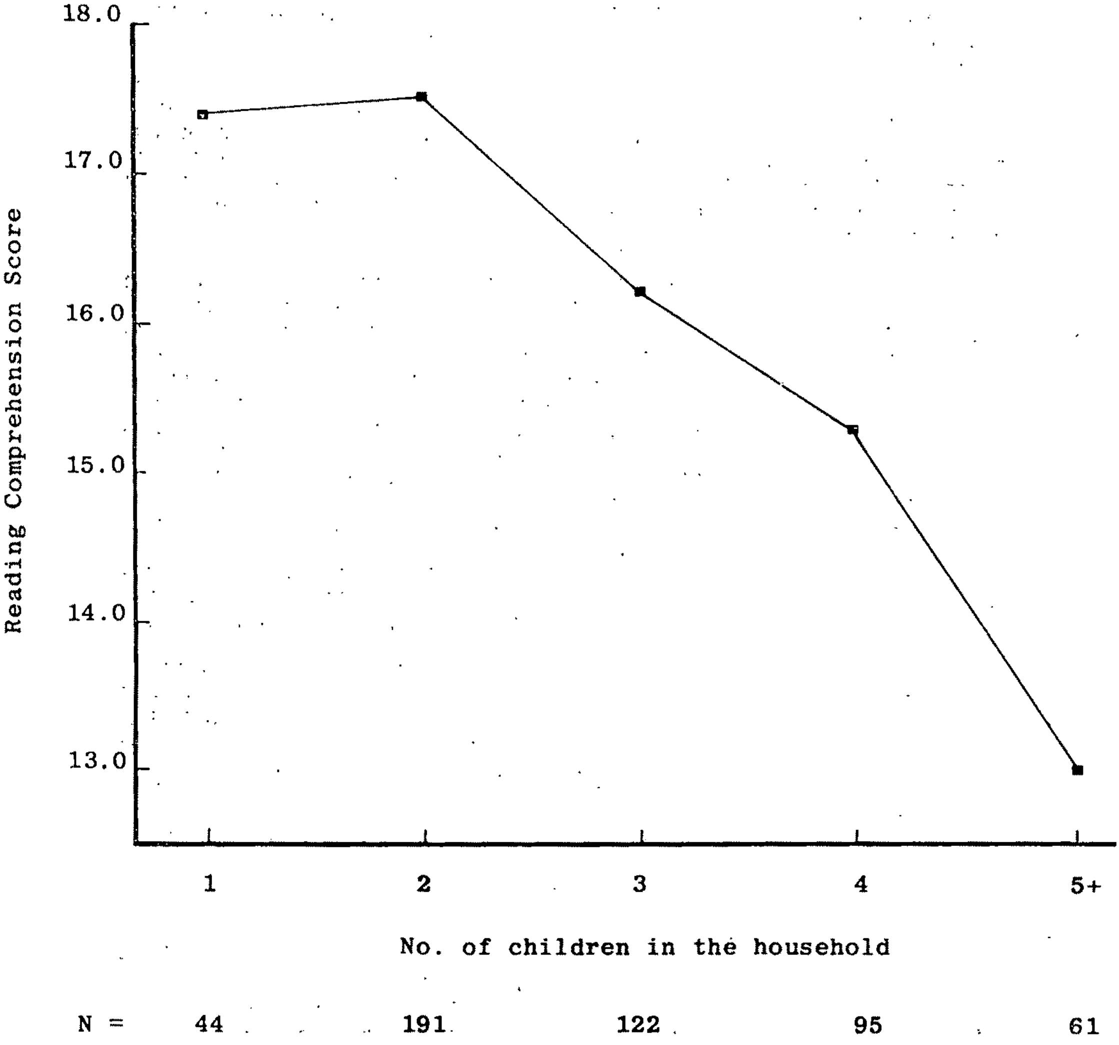
We concentrated on two aspects of interest in the analysis of these compositions. First, it seemed important to know the volume of children's writing; we thus used ' composition length in words ' as a measure of writing productivity. Second, indication of the syntactic richness, or structural complexity of the language written was wanted. Syntax is a major aspect of linguistic development; for example, it has been shown how school children use progressively more sentence-combining transformations in their speech and writing (Hunt, 1970). Moreover, this trend would seem to be part of a more general cognitive capacity, namely, " the ability to pack a greater density of ideas into a single sentence by embedding one sentence in another " (Cazden, 1972). A major problem, however, is how this ability can be represented meaningfully and quantitatively. We have adopted the T-unit devised by Hunt (1970). The T-unit, which is the length in words of units consisting of each main clause together with attached subordinate clauses and non-clausal structures, inevitably conceals a lot of interesting information. However, it has the advantage of exhibiting an increase in length as a direct function of syntactic complexity (Hunt, 1970; for full details of data collection and analysis, see Richardson *et al.*, 1976). Information regarding family size was obtained from parents in the course of an interview carried out in the home by health visitors. It was coded as the number of children under the age of 21 normally living in the household.

RESULTS

For some children not all details of the data were available so that the numbers in each family size group vary slightly according to which variable is being considered. Figure 1 shows a plot of reading comprehension score against number of children in the household for this sub-sample. The plot illustrates the typical average decline in test performance with increasing family size of the study child's home. This relationship was statistically highly significant. When the difference between family-size groups on composition length were examined, however, those were not found to be significant (Fig. 2). Furthermore, there were no significant family-size differences on use of syntactic resources as measured by mean T-unit length. These data are plotted in Figure 3.

FIGURE 1

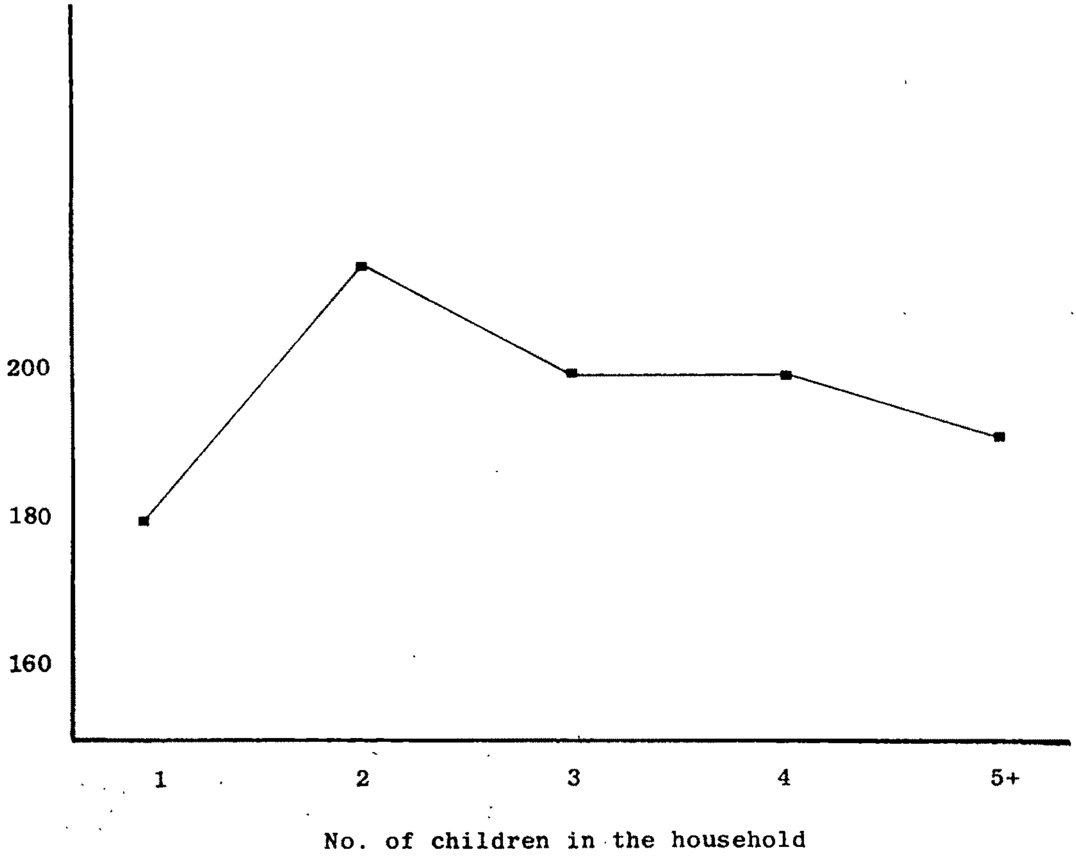
PLOT OF READING TEST SCORE AGAINST FAMILY-SIZE.



Analysis of variance: $\chi^2=43\cdot5$; $P<0\cdot001$.

FIGURE 2

Plot of Composition Length Against Family-size.

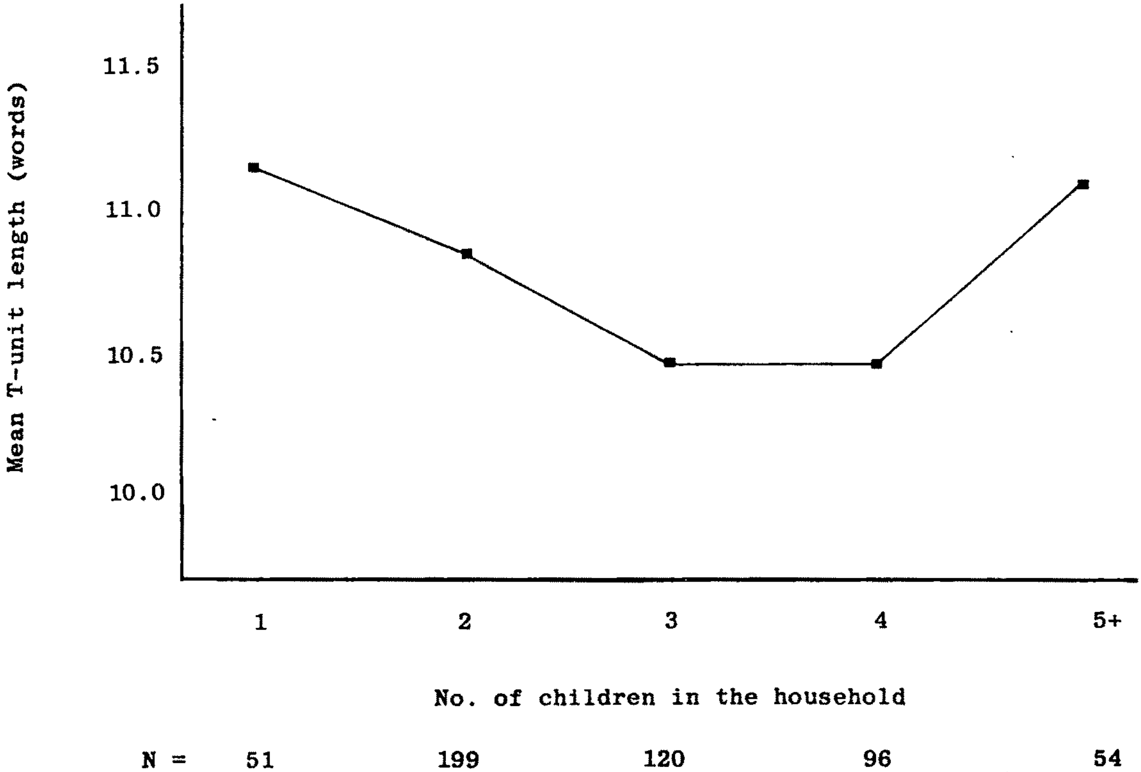


N =	51	199	120	96	54

Analysis of variance: $\chi^2=5\cdot6$; NS.

FIGURE 3

Plot of Mean T-unit Length Against Family-size.



Analysis of variance: $\chi^2=6{\cdot}7$; NS.

DISCUSSION

The anomaly emerging in the analysis described is that children from larger families (on average ' poorer readers ') perform as well as those from smaller families (on average ' better readers ') in terms of writing productivity and, in doing so, exhibiting an equivalent level of syntactic maturity, or sentence complexity, at least as measured by T-unit length. The usual explanation for the verbal attainment family-size (negative) correlations is that children from larger families receive relatively less parental verbal stimulus and/or suffer under inauspicious parental attitudes (Nisbet and Entwistle, 1967; Prosser, 1973).

Bormuth (1966) found over 60 factors contributing towards comprehension difficulty in reading test performance. Among these, syntactic elements such as sentence complexity or number of subordinate clauses had a substantial loading. Word difficulty and sentence length in themselves do not always seem to be significant aspects of passage readability among school children (for a review see Pearson, 1974). However, the role of syntax in constraining passage readability has been established in a number of studies (e.g., Epstein, 1961; Miller, 1962; Siler, 1973).

It should be noted that equality of T-unit length in the present sample does not preclude their being composed of different syntactic structures. Syntactic divergency associated with social class, for example, is now well established (Harms, 1963; Callary, 1971, 1974; Labov, 1969). The most prominent interpretations of such divergency are those based on a Whorfian perspective (e.g., Bernstein, 1972) which posit: (i) that speech systems are generated and regulated by forms of social relations, and (ii) that context is a major control upon syntactic usage. In so far as the Whorfian hypothesis has a level of generality, it might be expected that differentiation of syntactic usage might occur on the basis of family size; for example, it seems quite likely that families of different size will tend to generate different kinds of social relations. It has been noted that the reading performance/family size correlation holds within social class, and that of a number of measures examined family size accounts for by far the largest proportion of total variance in reading performance among 11-year-old children (Fogelman, 1975). In 7-year-olds this relative proportion is very much smaller with a word-recognition test (Davie *et al.*, 1972). Perhaps, therefore, the syntactic idiosyncrasies of (i) the reading test items and (ii) children from larger families, will help account for the relatively poor reading test performance of the latter, in spite of their apparent writing productivity and syntactic maturity. It would be interesting to know if this sort of explanation might be generalised to other studies using other tests, and if, in consequence, it has implications for reading materials in schools as well as selection of reading test items. Verifying it would, of course, demand a relatively thorough syntactic analysis of the language production of children from different sized families, perhaps of the sort used by Callary (1971) or, on the basis of T-units, by Peltz (1973).

ACKNOWLEDGMENTS—The author is grateful to M. Calnan, J. Essen and L. Lambert for marking of scripts, to H. Goldstein for carrying out the statistical tests, and to other colleagues of the National Children's Bureau for helpful comments and criticisms. This research was supported by grants from the DES and DHSS which are gratefully acknowledged.

REFERENCES

BERNSTEIN, B. (1972). A sociolinguistic approach to socialisation; with some reference to educability. In GUMPEZ, J. J., and HYMES, D. (Eds.), *Directions in Sociolinguistics.* New York: Holt, Rinehart and Winston.

BORMOUTH, J. R. (1966). Readability: a new approach. *Reading Res. Quart.*, 1, 79-132.

CALLARY, R. (1971). Syntactic correlates of social stratification. Unpublished doctoral dissertation, Louisiana State University.

CALLARY, R. (1974). Status perception through syntax. *Language and Speech*, 17, 187-192.

CAZDEN, C. B. (1972). *Child Language and Education.* New York: Holt, Rinehart and Winston.

DAVIE, R., BUTLER, N. R., and GOLDSTEIN, H. (1972). *From Birth to Seven.* London: Longman.

DAVIE, R. (1973). Eleven years of childhood. *Stat. News*, 33, 1-7.

DOUGLAS, J. W. B. (1964). *The Home and the School.* London: Fontana.

EPSTEIN, W. (1961). The influence of syntactic structures on learning. *Am. J. Psychol.*, 74, 80-85.

FOGELMAN, K. R (1975). Developmental correlates of family size. *Br. J. Soc. Wk.*, 5, 43-57.

FOGELMAN, K. R., and GOLDSTEIN, H (1976). Factors associated with changes in school attainments between seven and eleven. *J. educ. Res.*, 2, 95-109.

HARMS, L. S. (1963). Status, cues in speech: extra-race and extra-region identification. *Lingua*, 12, 300-306.

HUNT, K. W. (1970). Syntactic maturity in school children and adults. *Monogr. Soc. Res. Child Dev.*, 35 (i), Ser. No. 134.

LABOV, W. (1969). The logic of non-standard English. *Georgetown Monogr. Lang. Ling.*, 22, 1-31.

MILLER, S. A. (1962). Some psychological studies of grammar. *Am. Psychol.*, 11, 748-762.

NISBET, J. D., and ENTWISTLE, N. (1967). Intelligence and family size 1949-64. *Br. J. Educ. Psychol.*, 37 (2), 188.

PEARSON, P. D. (1974). The effects of grammatical complexity on children's comprehension, recall and conception of certain semantic relations. *Reading Res. Quart.*, 10, 155-192.

PELTZ, F. K. (1973). The effect upon comprehension of repatterning based on students' writing patterns. *Reading Res. Quart.*, 9, 603-621.

PROSSER, H. (1973). Family size and children's development. *Health Soc. Serv. Jl.* March 10.

RICHARDSON, K., CALNAN, M., ESSEN, J., and LAMBERT, L. (1976). The linguistic maturity of 11-year-olds ; some analysis of the written compositions of children in the National Child Development Study. *J. Child Lang.*, 3, 99-115.

SILER, E. R. (1973). The effects of syntactic and semantic constraints on the oral reading performance of second and fourth graders. *Reading Res. Quart.*, 9, 583-602.

(*Manuscript received 24th March*, 1976)

Br. J. educ. Psychol., 47, 76-80, 1977

FURTHER INVESTIGATION OF THE EMERGENCE OF ABILITY FOR PARTIAL AND COMPLETE MULTIPLICATION OF CLASSES IN YOUNG SCHOOL CHILDREN

By BERNICE Y. L. WONG

(*Faculty of Education, Simon Fraser University, Canada*)

SUMMARY. Piaget's theoretical position assumes that the ability to solve complete multiplication will be an earlier development in the child than partial multiplicative classification and his experimental findings support this hypothesis. The present experiment using 159 children aged 4:7 to 7:6 years re-examined the issue and investigated children's performance in these two types of multiplication using a consistent response mode to measure performance. The results substantiate Piaget's hypothesis. Further investigations were suggested to determine which of Piaget's two explanations is more applicable to the later development of partial multiplication.

INTRODUCTION

In their investigations of the developmental relation between complete multiplication and partial multiplication, Inhelder and Piaget (1964) found complete multiplication developed before partial multiplication. The authors proposed two reasons for this finding. According to Piaget's structuralist perspective (Piaget, 1970), the entire cognitive structure necessary for the function of an operation must exist before the child can successfully execute an operation. Because it is part of the operation of complete multiplication, partial multiplication would thus be a later development in the child. The other explanation for Piaget and Inhelder's finding concerns the facilitating effects of the perceptual symmetry in complete multiplication on the child's performance in this task.

One may legitimately question the validity of the interpretations and inferences presented above. Findlay (1971) has pointed out a serious flaw in Inhelder's and Piaget's (1964) experiment. The latter used inconsistent response modes to test children's performance in complete multiplication and partial multiplication. More specifically, in the complete multiplication experiment, the child had to choose a response from a finite number of alternatives. However, in the partial multiplication experiment, the child, in the majority of cases, had to make a creative response. Thus, response mode confounded Inhelder and Piaget's finding. Consequently, their obtained developmental sequence becomes suspect and a methodologically improved study is in order, to see if indeed one response mode facilitates children's performance more than the other.

Findlay (1971) gave children two partial multiplication tasks. For one group of subjects, the response mode involved finite choice from a set number of alternatives. For the other group of subjects, the response mode involved creative response. Findlay hypothesised that children would find the task easier in the finite choice situation and would make more correct responses than in the creative response situation. She reasoned that the probabilities of choice are limited in the finite situation, where the pictures are present. In the creative situation the choices for solution are not limited because the children are asked to construct their own responses. Findlay's hypothesis was substantiated. Her subjects found the finite response mode in partial multiplication was easier. (1) The children made significantly more correct responses in the finite response situation than in the creative response situation. (2) There was no improvement across age groups 6–10 in the creative response condition. (3) In contrast, there was significant improvement across the same age groups in the finite response condition. Thus, Findlay showed unequivocally that children's performance in partial multiplication was affected by the designated response mode. Her findings suggest that Piaget had inadvertently assigned a more difficult response mode to the partial multiplication task, and an easier response mode to the complete multiplication task. In sum, Piaget's obtained developmental sequence may well be an artefact induced by the methodological confounding of test modes.

While Findlay's study contributes to experimental methodology, that experiment did not deal directly with the question of whether children master complete multiplication before partial multiplication. The present experiment was designed to investigate this question, using a consistent response mode: the finite choice situation. This response measure was chosen because Findlay's study indicated that the creative response mode was insensitive in differentiating children's performance. The problem investigated in the present experiment is particularly important because of its implication for Inhelder and Piaget's earlier work, as well as for our understanding of children's cognitive development in multiplication of classes.

METHOD

Sample.

The sample consisted of 159 children aged 4 years 7 months to 7 years 6 months, drawn randomly from four nurseries, four kindergarten and five grade one classes from schools located in North Vancouver, B.C. Their parents' occupations were representative of those found in the North Vancouver population as a whole. Each subsample, nursery, kindergarten and grade one, contained 53 children. The ratio of boys to girls was approximately equal. The subjects were all Caucasian excepting for two Chinese siblings who spoke fluent English.

Procedure.

The materials for the complete multiplication task consisted of 15 picture cards. Each picture card measured 2 by 2½ inches (50·8 mm × 60·3 mm). The picture cards depicted geometric designs which could be grouped according to colour and form (circle, triangle and cross; yellow, red, and blue). Seated across the table from the subject, the experimenter presented that subject with a three by three matrix formboard made of cardboard, measuring 7½ by 6 inches (175·7 mm × 152 mm), and proceeded to fill all the cells consecutively except the diagonal ones. First, the experimenter filled cells 1 and 2 (see Figure 1) in the matrix and asked the subject to verbalise their contents, and to compare and contrast them. Then cell 3 was filled and the subject was asked to relate it to the first two cards in terms of their differences and similarities. Next cell 4 was filled and the whole procedure of comparison and contrast with the other cards repeated. Finally, cells 5 and 6 were filled and the same procedures of comparison and contrast with the rest of the cards was repeated.

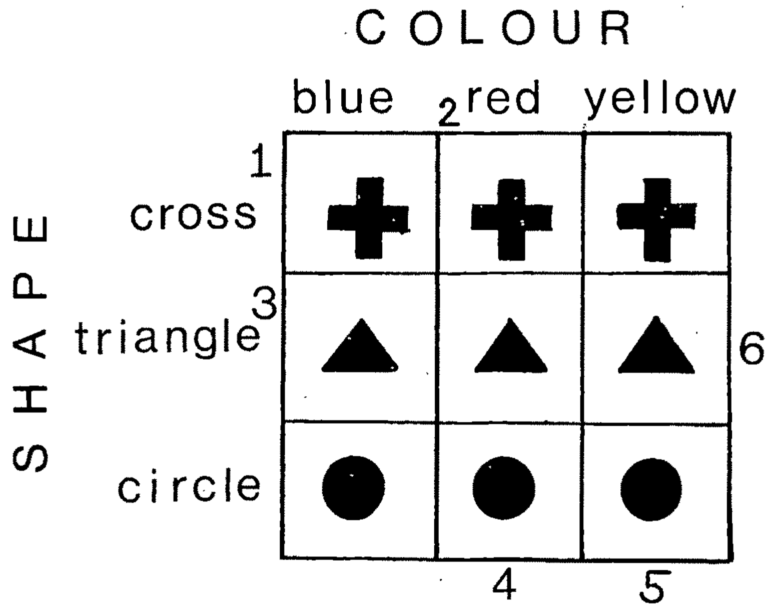
After the subject had filled an empty cell, the experimenter removed the two remaining unused choice cards and presented the subject new choice cards for the next empty cell. Only one of the three choice cards was the correct response card, the other two controlled for shape and colour. For example, for the middle diagonal cell, the correct response card was a red triangle. The remaining response cards were an orange triangle (control for shape) and a red square (control for colour). For the top right corner, the correct response card was a yellow cross, the incorrect cards were a purple cross (control for shape) and a yellow rectangle (control for colour). For the bottom left corner, the correct response card was a blue circle, the incorrect cards were a green circle (control for shape) and a blue hexagon (control for colour). The position of the correct choice card was randomised in each display of choice cards. The subjects received no feedback of right or wrong regarding their performance. For every correct response choice, the subject received one point. The scoring range was 0–3 points.

The partial multiplication task consisted of three subtasks. The first subtask replicated the version used by Inhelder and Piaget (1964). In this task, the subject was shown two vertical columns of pictures of green objects. The experimenter elicited from the subject the common quality here, that is, green colour. Next, the experimenter showed the two rows of pictures of leaves of diverse colours and elicited from the subject their common quality, i.e., leaf (shape). Then the experimenter put the stimuli in the form of a cross and asked the subject to fill in the space with a choice card out of five.

The second subtask consisted of drawings of flowers of different colours in the row and yellow objects in the column. The correct choice card was the picture of a yellow flower. The last subtask consisted of dogs drawn in black and white for the row and blue objects for the column. The correct choice card was a drawing of a blue dog. The last subtask was designed to see if young children would choose the logical response card even when in reality a blue dog does not exist. For each correct choice card, the subject received one point. The scoring range was 0–3 points. The reason for changing to pictures from symbols was to minimise perceptual confusion for the subjects.

FIGURE 1

STIMULI USED IN THE COMPLETE MULTIPLICATION TASK.



RESULTS

Using a criterion of two out of three correct responses as a pass and zero to one as failure, a frequency contingency table of pass/fail was constructed. The binomial test (Siegel, 1956) was used to test the developmental relations in Table 1. The binomial test was used instead of the chi-square test because it permits direct focus on the diagonal cells without using the marginal cells in testing for significance. The criterion of 2/3 was chosen rather than 3/3 because it was felt that the latter criterion may depress unduly the performance of the nursery children. The appropriateness in choosing 2/3 as a passing criterion is evidenced in the distribution of scores among subjects. In the complete multiplication task, 14 nursery subjects scored 2/3 and 5 scored 3/3, with the remaining 34 subjects scoring either zero or one. Twenty-one kindergarten subjects scored 2/3 and 9 scored 3/3, with 23 subjects scoring either zero or one. Fifteen grade-one subjects scored 2/3 and 23 scored 3/3, with the remaining 15 scoring either zero or one. In the partial multiplication task, 8 nursery subjects scored 2/3, 4 scored 3/3 with 41 scoring either zero or one. Seventeen kindergarten subjects scored 2/3, 7 scored 3/3 and 29 scored either zero or one. Twelve grade-one subjects scored 2/3, 21 scored 3/3 and 20 scored either zero or one. The pooled data of grade one, kindergarten and nursery children are summarised in Table 2.

The results indicated that children succeeded in the complete multiplication task significantly more often than in the partial multiplication task. This trend was consistently observed among the three groups of children studied: nursery, kindergarten and grade one. The ratios of children passing complete multiplication while failing partial multiplication were 15:6, 14:6, and 17:7 in the nursery, kindergarten and grade one groups, respectively (P< ·05; P< ·03; and P< ·001, respectively). The combined data from the three groups reflected the trend among the individual groups faithfully. The ratio of children mastering complete multiplication while failing partial multiplication and children evidencing the reverse performance sequence was 43:14. The difference was statistically significant (P< ·0001).

TABLE 1

THE PERFORMANCE OF NURSERY, KINDERGARTEN AND GRADE ONE CHILDREN ON COMPLETE AND PARTIAL MULTIPLICATION TASKS.

	Partial Multiplication of Classes								
School	Nursery			Kindergarten			Grade One		
Mean Age	4 years 7 months			5 years 6 months			7 years 6 months		
Complete Multiplication of Classes		Fail	Pass		Fail	Pass		Fail	Pass
	Pass	15**	4	Pass	14*	18	Pass	14***	25
	Fail	28	6**	Fail	15	6*	Fail	13	1***

* P< ·05 ** P< ·03 *** P< ·001

TABLE 2

THE PERFORMANCE OF 159 CHILDREN AGED 4 TO 7 ON COMPLETE AND PARTIAL MULTIPLICATION TASKS.

		Partial Multiplication of Classes	
		Fail	Pass
Complete Multiplication of Classes	Pass	43****	46
	Fail	56	14****

**** P< ·0001

DISCUSSION

The present experiment tested Inhelder and Piaget's (1964) hypothesised developmental priority of complete multiplication over partial multiplication, under a consistent response mode. The results substantiate Piaget's original hypothesis. One may contend that the obtained finding is confounded by the different content in the stimuli. The stimuli of the complete multiplication task consisted of simple geometric shapes while the stimuli of the partial multiplication task consisted of concrete shapes. The differential nature of the stimuli poses as a confound only if the geometric shapes used were highly unusual or unfamiliar. The shapes used in the present experiment are those typically seen in primary school materials, hence they cannot be considered unusual or unfamiliar. Moreover, similar geometric shapes were used by Inhelder and Piaget (1964, p. 160). The reason for not using geometric shapes in the partial multiplication task is that such a task is difficult to construct, without confusing younger children. For example, a conceivable version of it will have the column consisting of different geometric shapes of one colour while the row will consist of one geometric shape of different colours. One may contend that concrete shapes may have been used for the complete multiplication task. This has not been attempted in this experiment because a pilot study using cut-outs of turtles which varied in colour and orientation, showed marked ceiling effects among subjects aged 4 to 6.

Despite the support given to Piaget's hypothesis of the later development of partial multiplication, the present findings do not enable us to ascertain the precise underlying psychological processes. The two hypotheses presented by Piaget pertaining to the underlying psychological processes of the later development of partial multiplication appear equally applicable: (1) the facilitative functions of perceptual-figural cues in the symmetry of a matrix; and (2) the operational decalage in the child's abstracting portion of the whole structural system of complete multiplication. It is plausible that the matrix task format offers the child possible matching cues (standards) for his response choices. In case of doubt, he can always look at the rows and columns to check his response choice. The alternative theoretical interpretation also applies because of the feasible distinction between 'competence' and 'performance' (cf. Flavell and Wohlwill, 1969). The child may have developed the ability to perform a complete multiplication task. However, his newly developed ability may be suppressed by unfamiliar task format or stimuli. In short, for the full expression of his newly developed ability in multiplication of classes, the child needs to experience a variety of environmental stimuli. The present experimental design did not provide a test between Piaget's two explanations of the prior development of complete multiplication over partial multiplication. Future investigations to clarify this point are warranted.

REFERENCES

INHELDER, B., and PIAGET, J. (1964). *The Early Growth of Logic in the Child.* London: Routledge.

FINDLAY, A. D. (1971). An investigation of the emergence of ability for partial and complete multiplication of classes in young school children. *Br. J. educ. Psychol.*, 41, 96-98.

FLAVELL, J., and WOHLWILL, J. (1969). Formal and function aspects of cognitive development. In ELKIND, D., and FLAVELL, J. H. (Eds.), *Studies in Cognitive Development.* New York: Oxford Univ. Press.

PIAGET, J. (1970). *Structuralism.* New York : Basic Books.

SIEGEL, S. (1956). *Nonparametric Statistics for the Behavioral Sciences.* New York: McGraw-Hill.

(*Manuscript received* 3rd *May*, 1976)

Br. J. educ. Psychol., 47, 81-84, 1977

THE DEVELOPMENT OF CAUSAL AND LOGICAL CONNECTIVES IN CHILDREN

By J. T. LAWTON
(*University of Wisconsin, Madison, Wisconsin, USA*)

SUMMARY. Children's use of causal and logical connectives was examined before and after treatments based on advance organisers with social studies lessons. In samples of 120 children from age-levels 6 and 10 years a diminution in the use of syncretic reasoning and understanding followed the experimental treatment. The performance of the older group was generally better than that of the younger children, supporting Piaget's concept of stage-invariant cognitive development.

INTRODUCTION

This study arose from the testing of 6- and10-year-old children who had taken part in an expository teaching/reception learning intervention study, based on Ausubel's theory of meaningful reception learning, designed to enhance the development of subject-matter concepts, and concrete operations as described by Piaget. Piaget has stated that, if the child has a meaningful understanding of the integrated conceptual schema, within both the content he is handling and the operations he performs upon it, and has the ability to verbalise about conceptual relations, there should be an absence of juxtaposition and of syncretic reasoning and understanding (Piaget, 1926, p. 135). From his own data, Piaget concluded that, on average, causal explanation begins at about age 7 or 8 years and logical justification develops some time later (Piaget, 1928). Peel (1959, 1960) suggests that logical reasoning is unlikely to occur before age 13 years. Recently, Corrigan (1975) reported that the ability to use the concrete logical ' because ' marks the concrete operational stage. Ausubel (1963, 1969) has hypothesised that prior learned organisers (subsumers) in the form of abstract concepts and propositions influence the subsequent learning of particular new information when the relationship between ' subsumer ' and new knowledge is non-arbitrary. Ausubel and Schiff (1954) and Ausubel and Robinson (1969) suggest that children are restricted in the use of logical operations and the psychological processes involved in causal thinking by the measure of their capacity to relate relatively general ideas to specific instances. It was predicted that ' advance organiser ' lessons would facilitate both children's learning of meaningful verbal materials (high-level concepts and propositions) and logical operations (hierarchical classification). Such increments of learning were expected to lead to a diminution of syncretic reasoning and understanding evidenced by improved performance in the use of the causal and logical ' because.'

METHOD

Sample and materials.

Social studies learning materials were selected which described various communities of primitive and modern people (past and present) and how they satisfied certain basic needs with reference, mainly, to food, dwellings and tools. Advance organiser (AO) lessons presented this information as high-level concepts and propositions (illustrated by concrete examples). These lessons also presented high-order rules for hierarchical classification. For pre-and post-tests, a sentence completion task was used similar to that used by Piaget (1926). Unlike Piaget's procedure of selecting arbitrary sentence materials, the post-test sentences used in this study presented material from the social studies content referred to. For both pre-and post-tests five sentences required a causal connection to be established, and five a logical justification.

Subjects were 120 children selected at random in each of the age-groups 6/7 years and 10/11 years from two primary schools in an urban area of Northern England, so that each group was comprised of 60 children. The English Picture Vocabulary Tests I and II were administered and children were placed in two experimental groups and one control group for each group so that each age group was equally represented in terms of VR score.

Procedure.

Two related experiments were conducted. In each experiment an AO lesson was taught to the experimental groups. The first dealt with general ideas about stone-age and early-Eskimo people (primitive) and a present-day community (modern) along with high-order rules for hierarchical classification. The second AO lesson dealt with the communities of Bedhouin herdsmen, Thai rice-farmers (primitive) and Norwegian farmer-fishermen (modern) plus the same high-order rules for hierarchical classification as in the first AO. In addition, the first experimental group in both age groups received *subsequent* lessons (SL) providing further particular examples (low order concepts) related to the high-level concepts of the AO lessons. The control groups received *introductory* lessons and the *subsequent* lessons which provided only low-order concepts. For pre- and post-tests a short, illustrated story was read to each child individually describing events in the lives of the people referred to in the AOs. The story was followed by the sentence-completion items. The first part of each sentence was read by the experimenter, finishing, in each case, with ' because.' The child completed each sentence, encouraged to do so as well as possible. Pre-tests referred to communities of people not included in the AO lesson materials. Responses were scored correct or incorrect according to three criteria:

(1) each response must be correct within the context of the story and related material;

(2) responses to the first five items must give a correct explanation of a cause and effect relationship;

(3) responses of the second five items must show a relationship of implication, of reason and consequent.

RESULTS

Experiment 1.

The mean number of correct statements for each type of item are given in Table 1.

TABLE 1

MEAN NUMBER OF CHILDREN RESPONDING CORRECTLY TO ' CAUSAL ' AND ' LOGICAL ' ITEMS: EXPERIMENT 1.

		Pre-Test		Post-Test	
Age-treatment Groups	N	Causal Items	Logical Items	Causal Items	Logical Items
6—E1 (AO+SL)	20	0	0	13**	7**
6—E2 (AO)	20	0	0	9 NS	4 NS
6—Control	20	0	0	0·4 NS	0·2 NS
10—E1 (AO+SL)	20	0·6	0·4	18**	12**
10—E2 (AO)	20	1·2	0·6	11**	7**
10—Control	20	1·4	0·4	5*	1·6 NS

* P< ·05 ** P< ·01 between pre- and post-test results.

None of the 6 and few 10-year-olds completed the pre-test sentences correctly. Lack of scores, or low scores makes any statistical procedure unnecessary. All these children found both types of sentence-completion item difficult. In responding they often changed the syntax of the intended ' because ' clause to a ' for ' or ' then ' clause.

Nor were there many attempts to explain. Children tended to assume a fact was correct, not seeking to explain the phenomenon. They tended to leap to conclusions which are the outstanding characteristics of verbal syncretism. Table 1 indicates that in most cases experimental groups significantly improved their performance on the post-test.

The experimental groups performed significantly better than control groups on the post-test. This difference was more noticeable on causal 'because' items: 6-year-olds causal—items, $P<{\cdot}01$; logical items, $P<{\cdot}01$ (χ^2); 10-year-olds—causal items, $P<{\cdot}01$, logical items, $P<{\cdot}01$ (χ^2). The 10-year-old experimental group performance was significantly better than their younger counterparts (causal items, $P<{\cdot}01$; logical items, $P<{\cdot}05$, χ^2). The performance of the older control group children were superior to the 6-year-old control group, at least on causal items ($P<{\cdot}05$, χ^2). There was a strong indication, generally, that performance improved partly as a function of age.

Finally, there was a strong indication, in Experiment 1, that the combination of AO with SL was more effective than the AO on its own.

Experiment 2.

The experiment of children's further use of the causal and logical 'because' replicated the details of Experiment 1. The number of correct responses to sentence-completion items on Experiment 2 pre- and post-tests are summarised in Table 2. Lessons and task content referred to primitive and modern communities different from those in Experiment 1.

TABLE 2

MEAN NUMBER OF CHILDREN RESPONDING CORRECTLY TO 'CAUSAL' AND 'LOGICAL' ITEMS: EXPERIMENT 2.

		Pre-Test		Post-Test	
Age-treatment Groups	N	Causal Items	Logical Items	Causal Items	Logical Items
6—E1 (AO+SL)	20	13	9	17	12
6—E2 (AO)	20	12	10	16	13
6—Control	20	2	1	1	0·8
10—E1 (AO+SL)	20	17	15	18	17
10—E2 (AO)	20	15	13	15	14
10—Control	20	5	6	3	5

First, a comparison of Experiment 2 pre-test and Experiment 1 post-test results indicated a positive transfer effect which can be supposed to result due to the meaningful relationship between pre-test Experiment 2 items to the general ideas of the *first* AO lesson. Further test characteristics can be briefly reviewed. There was a general invariance on causal and logical items by experimental groups. However, a tendency for causal items to be more easily dealt with was apparent, especially by the 6-year-old experimental groups. Both experimental groups maintained a similar level of performance for pre- and post-tests. Their performance was superior to that of control groups ($P<{\cdot}01$, χ^2).

DISCUSSION AND CONCLUSIONS

It was found that initial responses indicated both age-groups of children were experiencing considerable difficulty with causal explanation and logical justification. These results were not expected in the case of 10-year-olds and task content appeared to be having a crucial effect on performance. The expository teaching procedure proved effective for experimental groups regardless of age. The crucial learning effect appears to be an improvement in the use of causal explanation and logical justification resulting from a utilisation of high-order concepts and propositions that have the power to explain special relationships between more particular instances. Thus, Piaget's conclusions regarding syncretism are not supported. However, at the same time, in spite of a likely task ceiling effect, there was a strong tendency for performance to improve with age, thus supporting Piaget's concept of ' stage ' sequence. Naivety as regards task content or experience in dealing with this type of task seems to have a considerable effect on performance. As the two experiments progressed data indicated that the meaningful reception learning of high-order concepts produced an acceleration of learning for the 6-year-olds. The improved performance of the 10-year-old experimental group children can be described in terms of enhancement rather than acceleration. Thus, in both cases Ausubel's hypothesis was supported. Bruner (1960, p. 369) suggests that the concrete operational child need not manipulate the world of potentiality until pressure is put upon him which results in the beginning of propositionalism. The acceleration and enhancement of performance, resulting from the AO lessons, not only supports Ausubel's proposition that stable structures of ideas can be meaningfully related to particular instances but that meaningful reception learning appeared to provide, in this case, a *demand* which brought about a qualitative change in thinking.

These results must be treated with circumspection. This study of necessity, was intended to be exploratory in nature.

ACKNOWLEDGMENTS—The data for this study were collected as part of a larger investigation submitted in partial fulfilment of the requirements of the Ph.D. degree at Leeds University, England. Thanks are due to Professor K. Lovell for his detailed guidance and advice. The author, however, is entirely responsible for the views expressed.

REFERENCES

AUSUBEL, D. P. (1963). *The Psychology of Meaningful Verbal Learning.* New York: Grune and Stratton.

AUSUBEL, D. P. (1969). *Readings in School Learning.* New York: Holt, Rinehart and Winston.

AUSUBEL, D. P., and SCHIFF, H. M. (1954). The effect of incidental and experimentally induced experience in the learning of relevant and irelevant causal relationships by children. *J. genet. Psychol.*. 84, 109-129.

AUSUBEL, D. P., and ROBINSON, F. G. (1969). *School Learning: An Introduction to Educational Psychology.* New York: Holt, Rinehart and Winston.

BRUNER, J. S. (1960). *The Process of Education.* Cambridge, Massachusetts: Harvard University Press.

CORRIGAN, R. A. (1975). Scalogram analysis of the development of the use and comprehension of ' because ' in children. *Child Dev.* 46 195-201.

PEEL, E. A. (1959). Experimental examination of some of Piaget's schemata concerning children's perception and thinking and a discussion of their educational significance. *Br. J. educ. Psychol.* 29 89-103.

PEEL, E. A. (1960). *The Pupil's Thinking.* London: Oldbourne Press.

PIAGET, J. (1926). *The Language and Thought of the Child.* London: Routledge and Kegan Paul.

PIAGET, J. (1928). *Judgment and Reasoning in the Child.* London: Routledge and Kegan Paul.

(*Manuscript received 22nd March*, 1976)

Br. J. educ. Psychol., 47, 85-90, 1977

THE WPPSI AT FOUR YEARS: A SEX DIFFERENCE IN VERBAL-PERFORMANCE DISCREPANCIES

By H. FAIRWEATHER and G. BUTTERWORTH
(*Department of Experimental Psychology, University of Oxford*)*

SUMMARY. 72 boys and 72 girls at age 4½ years were tested on the WPPSI. Boys outscored girls significantly on Mazes. Twenty per cent of the children had discrepancies between the Verbal and Performance sub-scales of 15 points or more. Low verbal and low performance sub-groups were equally represented across sexes; but their location with respect to normative sub-scale scores was not. Boys' sub-groups were distinguished by changes in performance scores alone whereas for girls both performance and verbal quotients deviated equally. The results underline the necessity for comprehensive control data on materials used in studies of cognitive development.

INTRODUCTION

The most appropriate individual intelligence test for children of below 5 is the Wechsler Pre-School and Primary Scale of Intelligence (WPPSI), which has thus far been standardised on two samples of British school children at age 5½ (Brittain, 1969; Yule *et al.*, 1969). Both studies find that boys generally outscore girls, but only three subtest differences reach statistical significance: girls on Animal House (this being the only subtest in either study in which girls substantially outscore boys), and boys on Mazes (Yule *et al.*) and boys on Vocabulary in the smaller groups used by Brittain.

Although often felt to be unstable below 6 years (Beldoch, 1963), the WISC itself had been widely used with 5-year-olds prior to the advent of the WPPSI. At this age, the original American norms show slight advantages to girls on the Performance sub-scale, but little difference on Verbal subtests. The pattern of such differences is not consistent, however, until 10 years when boys begin to dominate on both (Seashore *et al.*, 1950). From the same data, Miele (1958) notes only two subtests yielding sex differences: Digit Symbol Substitution or Coding (favouring girls) and Mazes (favouring boys), the latter being a supplementary test not included in the WISC assessment of Performance IQ. In a separate sample, Darley and Winitz (1961) find girls significantly ahead again on Coding, and also on Similarities and Performance IQ.

Verbal Performance discrepancies on these tests are the rule rather than the exception: in over half of all children tested they amount to eight points or more (Seashore, 1951; Brittain, 1969; Yule *et al.*, 1969). As Seashore has pointed out, such discrepancies should be interpreted with caution, particularly since they may well fall within the standard error of the sub-scales. The standard errors on the WPPSI at 4½ years are 3·63 points (Verbal) and 4·02 (Performance) which gives a possible error for difference of over 15 points. A difference of this magnitude, though, " becomes important and should be investigated " (Wechsler, 1967). Even then it is likely that one would be tapping a quarter of the population (Seashore, 1951; Brittain, 1969; Yule *et al.*, 1969). Within such a group, Yule *et al.* found a large majority of children with lower Verbal IQ. However, this finding was not detailed with respect to sex, or to the actual sub-scale scores. The present study was designed to explore sex differences in these discrepancies in a younger population than those already tested.

METHOD

Sample. The study formed part of a pre-school assessment programme carried out with children from a rural area. All the children aged 4½ years, registered with a general practice in the period April, 1972, to April, 1974, were tested. The social stratification of the sample reflected that of the general population of the area, which is slightly under-represented in Social Classes I and II. As there are only two general medical practices in the town, there is no reason to suppose any serious bias in the sample.

---

* The authors are now at Istituto di Psicologia dell'Universita di Bologna (Fairweather) and Department of Psychology, University of Southampton (Butterworth).

Procedure. Parents were asked to bring their child to the surgery where testing took place in a quiet room. All the children were tested by the same examiner (GB). Most of the children co-operated well. However, a small minority refused to co-operate at all. Although some of these children were subsequently tested at home, they have been left out of the analysis.

The children were seen within one month of reaching the age 4 years 6 months.

In all, 181 children (85 boys, 96 girls) were seen. Of these one boy and eight girls were too shy or otherwise unco-operative to complete a single subtest. A further 11 boys and 12 girls refused to do at least one subtest, and one boy had an IQ too low (<70) to be tested. Four girls were excluded randomly to leave balanced groups (N=72).

RESULTS

Normative data on the WPPSI. The means and standard deviations for each subtest, and the Verbal, Performance and Full Scale scores are listed by sex in Table 1. Sex differences were investigated by a series of one-way analyses of variance. Of the 13 such tests, one only reached significance. This was boys' superiority on Mazes (P< ·001). Boys tend also to score higher on Information, Block Design and Performance IQ; and girls on Comprehension, and Geometric Design.

TABLE 1

THE WPPSI: NORMATIVE DATA AT FOUR-AND-A-HALF YEARS.

Sub-tests	Boys (N=72)		Girls (N=72)		Significance of sex differences	
	Mean	SD	Mean	SD	F Ratio	P
Information....	9·80	2·37	9·35	2·75	1·15	NS
Vocabulary	9·80	2·11	9·86	2·07	<1	
Arithmetic	8·96	2·15	8·86	2·48	<1	
Similarities	10·92	2·49	11·16	2·46	<1	
Comprehension	9·86	2·09	10·38	2·46	1·91	NS
Verbal:........	99·39	10·37	99·38	11·11	<1	
Animal House ..	9·61	2·32	9·71	2·66	<1	
Picture Compln.	10·07	2·73	9·78	2·47	<1	
Mazes	10·46	2·78	8·92	2·59	11·85	< ·001
Geom. Design ..	9·64	3·09	10·35	2·27	2·48	NS
Block Design ..	10·68	3·08	9·92	2·27	2·87	NS
Performance: ..	100·60	14·27	98·24	11·40	1·20	NS
IQ	100·68	12·10	98·64	10·95	<1	

Verbal-Performance discrepancies. The numbers of children with discrepancies in the ranges 10–14 points, and 15 points and above, are shown in Table 2. 40 per cent of all children had discrepancies of over 10 points, and 20 per cent of 15 points or above.

TABLE 2

NUMBER OF CHILDREN WITH LARGE VERBAL PERFORMANCE DISCREPANCIES.

Group	V>P		P>V	
	15+	10–14	10–14	15+
Boys	6	3	5	9
Girls	8	12	9	6

Whilst there appears to be a tendency for more girls to show discrepancies in the 10–15 range, this is not significant ($\chi^2=0{\cdot}40$, NS). There are clearly no sex differences in frequency at 15 and above, nor any tendency for these discrepancies to be predominantly P>V, or V>P.

Sub-scale scores for the High discrepancy groups ($V\sim P\geqslant 15$ points) are detailed by subtests in Table 3, together with the significance levels of the relevant statistical comparisons (by t-test) both within and between sexes.

TABLE 3

MEAN SUBSCALE AND SUBTEST SCORES FOR HIGH DISCREPANCY GROUPS ($V\sim P\geqslant 15$ POINTS).

Sub-Scales	Boys		Girls		Level of Significance t-Test, P<			
	V>P	P>V	V>P	P>V	Within-sex		Across-sex	
					Boys	Girls	V>P	P>V
Information	9·3	10·0	11·1	8·3	NS	·10	NS	NS
Vocabulary.............	10·0	9·6	10·6	10·3	NS	NS	NS	NS
Arithmetic	7·7	9·6	11·5	6·8	NS	·01	·02	·001
Similarities	11·8	9·4	12·4	9·3	NS	·10	NS	NS
Comprehension	10·8	8·7	12·1	9·2	·10	·05	NS	NS
Verbal:	99·5	96·4	109·9	92·3	NS	·10	·10	NS
Animal House	7·0	11·8	9·0	10·5	·01	NS	NS	NS
Picture Completion	7·0	11·8	9·3	11·5	·001	NS	NS	NS
Mazes	8·2	13·6	7·5	12·2	·01	·01	NS	NS
Geometrical Design	6·7	12·7	9·9	11·7	·001	·10	·02	NS
Block Design	7·0	12·8	8·3	11·2	·01	·05	NS	NS
Performance:	80·5	117·1	91·8	109·8	·001	·05	·10	NS

For the boys V>P group, VIQ is almost identical with that of the normative sample (cf. Table 1) whereas PIQ is 20 points below the norm, with all subtests down by at least 2 points, and Block Design by over 3½. The VIQ for this group is also not significantly different from that of the P>V boys group, and only for Comprehension among the subtests does a difference approach significance (P< ·10). On the other hand, P>V boys' PIQ is some 16½ points higher than the norm, and consequently heavily outscores that of the V>P group (P< ·001), this disparity being reflected on every subtest.

In contrast, findings for the girls are much more even. In each discrepant group, the higher sub-scale score is about 10 points higher than the norm, and the lower some 6–7 points below. Correspondingly, the high scorers outscore the low scorers within each sub-scale to similar extents (P< ·10 for Verbal; P< ·05 for Performance).

Across-sex comparisons fail to significantly differentiate P<V groups, with the solitary exception of Arithmetic among Verbal subtests, and although boys outscore girls on all Performance subtests. Amongst V<P groups, girls tend to outscore boys on both VIQ (P<·10; and particularly for Arithmetic, P<·02), and PIQ (P<·10; and particularly on Geometric Design, P<·02).

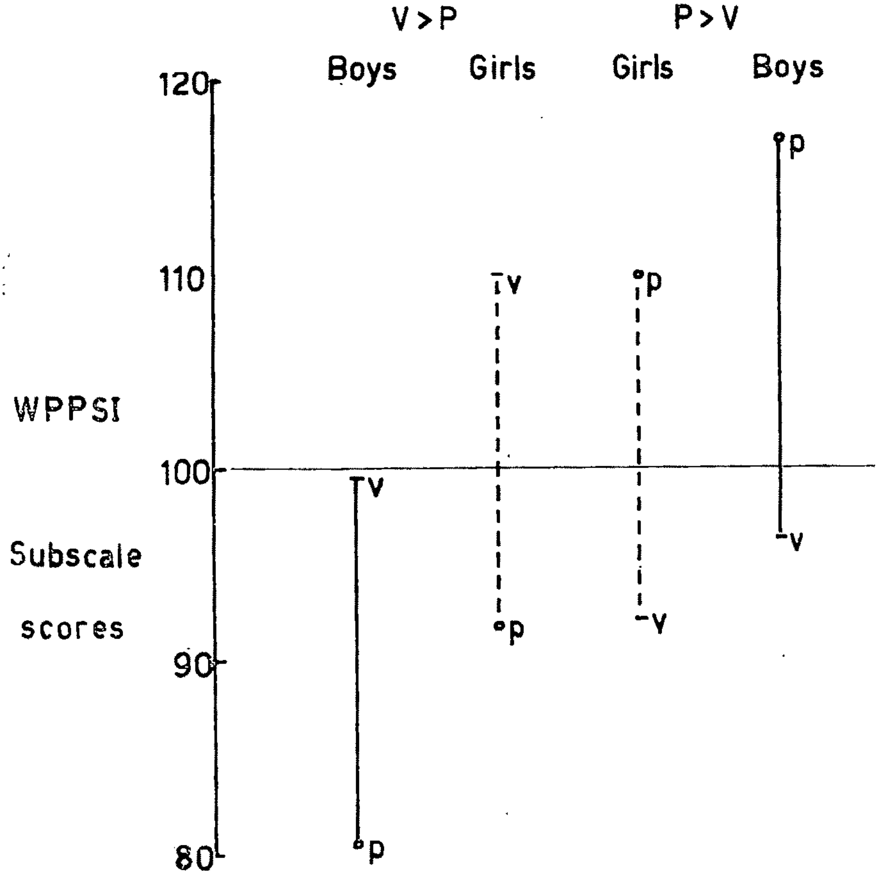
TABLE 4

COMPARISON OF HIGH *versus* LOW SCORERS ACROSS DISCREPANCY GROUPS.

Discrepancy Groups		t	P<
Boys	Highs ...... Lows	2·67 2·52	·02 ·05
Girls	Highs ...... Lows	0·06 0·09	NS NS

FIGURE 1

LOCATION OF VERBAL~PERFORMANCE DISCREPANCIES ≥ 15 POINTS ON THE WPPSI AT FOUR-AND-A-HALF YEARS.



The nature of the sex difference most clearly emerges from comparisons *across* sub-scales (see Table 4 and Figure 1). Amongst girls neither high nor low scores are significantly differentiated, whereas for boys they are: PIQ for P>V is greater than VIQ for V>P ($P < \cdot 02$); and VIQ for P>V is greater than PIQ for V>P ($P < \cdot 05$). The pattern is clear: whilst for girls both PIQ and VIQ deviate equally from the norm in generating a discrepancy, boys in contrast either promote or relegate only their PIQ, leaving VIQ ' anchored ' at the normative level.

DISCUSSION

In the more general aspects of performance level and sex differences the present data compare well with the earlier studies of British children at 5½ years (Brittain, 1969; Yule *et al.*, 1969). The only reversed sex difference with respect to both previous studies is girls' non-significant superiority on Geometric Design. Yule *et al.*, on the other hand, are alone in finding girls significantly better on Animal House. This is not altogether surprising since this test is the counterpart of WISC Coding which so consistently favours girls. Whilst some might argue that this reflects an earlier development of verbal labelling strategies, it would hardly be consonant with the current findings on Verbal IQ, or indeed with other indices of verbal competence at this age (e.g., Mittler and Ward, 1970). It is more likely to be a reflection of the changed motor component: Animal House requires arm/hand movements, whereas responses on Coding require the finer musculature involved in writing. Writing speed is not only much faster in girls, but has been directly related to the sex difference on WISC Coding (Lyle and Johnson, 1974).

Elsewhere on Performance subtests there is good agreement: no differences on Picture Completion, a small advantage to boys on Block Design, and a convincing superiority for boys on Mazes. This last is almost universally attested-to (see Wilson, 1975), though it is interesting to note that similar sex differences at the same age have disappeared with practice, rendering appeals to fixed biological substrata for this and related ' spatial ' tasks a little premature (McGinnis, 1929).

Together with Yule *et al.* (1969) for the WPPSI, and Miele (1958) and Darley and Winitz (1961) on the WISC, we find little to choose between the sexes on Verbal IQ. Brittain (1969) finds his girls relatively poor verbally (significantly so in Vocabulary): Darley and Winitz, on the other hand, produce a significant female superiority on Similarities. The Illinois Test of Psycholinguistic Abilities (ITPA) gives a small overall advantage to girls at 4½ years, but only one subtest reaches significance (Mittler and Ward, 1970).

The incidence of high Verbal-Performance discrepancies matches almost exactly those of earlier studies, but whereas Yule *et al.* (1969) report a predominance of P>V individuals (15·3 per cent, against 5·4 per cent for V>P cases, at 16 points and above) we find almost equal numbers in the two categories. The fact that Yule's sample was " over-representative of Social Classes I and II by about 9 per cent and under-representative of Social Class III (manual) by a similar amount " (p. 5) does little to explain this since their class distribution ought to increase the number of high discrepancies in favour of Verbal Scores. The interaction between class and sex in Yule's paper appears intricate: superior class elevates all sub-scales with the exception of boys' Verbal. In specific contradistinction to the present study, Yule *et al.* also tested children who were already at school and, therefore, pre-selected for a certain degree of verbal ability. This ought additionally to discriminate against the inclusion of low Verbal boys.

The present study extends the evidence on high VIQ–PIQ discrepancies in early childhood by identifying the locations of the two types of discrepancy with respect to the sub-scale norms for both sexes. Not only is the *balance* of importance, but also the *baseline* against which it is expressed. Notably, there are groups of boys with either very high or very low performance IQs. Even at the pre-school stage there is no evidence of a significant number of boys with peculiarly low or high verbal IQs. The present study gives no evidence of a significant number of girls with peculiarly high or low ability in either mode.

In conclusion, this study emphasises the need to report sex differences as well as grouped data when establishing baseline measures for tests commonly used in both normal (Fagan-Dubin, 1974) and clinical studies of cognitive development (Annett *et al.*, 1961; Fedio and Mirsky, 1969; McFie, 1961; Rudel *et al.*, 1974).

ACKNOWLEDGMENTS—The data reported here were obtained as part of a pre-school assessment programme carried out by the second author and financed by the Department of Health and Social Security in a grant to Dr. A. M. Steele, The Surgery, Chipping Norton, Oxon., England. The authors are indebted to Dr. Steele for permission to report the data in this form. The writing of the report was financed in part by grants to the first author from the Social Science Research Council of Great Britain, and the Royal Society and from the Medical Research Council to the second author.

REFERENCES

ANNETT, M., LEE D., and OUNSTED, C. (1961). *Intellectual Disabilities in Relation to Lateralised Features of the EEG.* London: Little Club Clinics in Developmental Medicine No. 4, Spastics Society/Heinemann.

BELDOCH, M. (1963). Applicability of the norms of the Wechsler Intelligence Scale for Children to 5-year-olds. *J. consult. Psychol.* 27, 263-264.

BRITTAIN, M. (1969). The WPPSI: a Midlands study. *Br. J. educ. Psychol.*, 39 14-17.

DARLEY, F. L., and WINITZ, H. (1961). Comparison of male and female kindergarten children on the WISC. *J. genet. Psychol.*, 99, 41-49.

FAGAN-DUBIN, L. (1974). Lateral dominance and development of cerebral specialisation. *Cortex*, 10, 69-74.

FEDIO, P., and MIRSKY, A. F. (1969). Selective intellectual deficits in children with temporal lobe or centrencephalic epilepsy. *Neuropsychologia*, 7, 287-300.

LYLE, J. G., and JOHNSON, E. G. (1974). Analysis of WISC coding. Prediction of coding performance. *Percept. mot. Skills*, 39, 111-114.

MCFIE, J. (1961). Intellectual impairment in children with localised post-infantile cerebral lesions. *J. Neurol. Neurosurg. Psychiat.*, 24, 361-365.

MCGINNIS, E. (1929). The acquisition and interference of motor habits in young children. *Genet. Psychol. Monogr.*, 6, 203-309.

MIELE, J. A. (1958). Sex differences in intelligence: the relationship of sex to intelligence as measured by the Wechsler Adult Intelligence Scale and the Wechsler Intelligence Scale for Children. *Diss. Abs.*, 18, 2213.

MITTLER, P., and WARD, J. (1970). The use of the Illinois Test of Psycholinguistic Abilities on British 4-year-old children: a normative and factorial study. *Br. J. educ. Psychol.*, 40, 43-54.

RUDEL, R. G., TEUBER, H. L., and TWITCHELL, T. E. (1974). Levels of impairment of sensorimotor functions in children with early brain damage. *Neuropsychologia*, 12, 95-108.

SEASHORE, H. G. (1951). Differences between verbal and performance IQs on the Wechsler Intelligence Scale for Children. *J. consult. Psychol.*, 15, 62-67.

SEASHORE, H., WESMAN, A., and DOPPELT, J. (1950). The standardisation of the Weschler Intelligence Scale for Children. *J. consult. Psychol.*, 14, 99-110.

WECHSLER, D. (1967). *Manual for the Wechsler Pre-school and Primary Scale of Intelligence.* New York: Psychological Corporation.

WILSON, R. S. (1975). Twins, patterns of cognitive development as measured on the Wechsler Pre-School and Primary Scale of Intelligence. *Dev. Psychol.*, 11, 126-134.

YULE, W., BERGER, M., BUTLER, S., NEWHAM, V., and TIZARD, J. (1969). The WPPSI: an empirical evaluation with a British sample. *Br. J. educ. Psychol.*, 39, 1-13.

(*Manuscript received* 13*th April*, 1976)

Br. J. educ. Psychol., 47, 91-94, 1977

A NOTE ON JENSEN'S THEORY OF LEVEL I ABILITY AND RECENT RESEARCH ON HUMAN MEMORY

By M. J. LAWSON
(*School of Education, Flinders University, South Australia*)
AND R. F. JARMAN
(*Department of Educational Psychology, University of British Columbia, Canada*)

SUMMARY. The assumptions of Jensen's theory of Level I ability are examined in the light of recent research on memory. Experimental studies of organisation in both serial and free recall, as well as factor analytic investigations of sequential memory indicate that the current theory requires revision in order to accommodate recent findings. Any revision should include the effects of subject-initiated learning strategies, and give more consideration to the nature of transforming processes.

INTRODUCTION

The two level theory of mental abilities proposed by Jensen (1968, 1969, 1970a) has been advanced as an explanation of relationships between such variables as intelligence, learning ability, and socio-economic status, and has formed the basis of a taxonomy of types of mental retardation (Jensen, 1970b). While the range of phenomena investigated within this framework is impressive, recent research has suggested the need for a re-examination of the assumptions incorporated in the two level theory. It is argued here that the distinction between the two ability levels is inadequate, and specifically, that the definition of Level I ability is not consistent with results in other independent research.

Central to the Levels' theory is the notion of transformation of informational input. Jensen proposes that the nature of the processing carried out in different mental tasks can be represented in terms of points on a continuum of transformation. The degree of transformation is reflected empirically in the amount of correspondence between input and output. The end points of this continuum define two qualitatively different types of abilities, Level I (associative) ability, and Level II (cognitive) ability. The differences between the ability levels are described by Jensen as follows :

> " Level I ability is essentially the capacity to receive or register stimuli, to store them, and to later recognise or recall the material with a high degree of fidelity . . . It is characterised especially by the lack of any need of elaboration, transformation or manipulation of the input to arrive at the output . . . Level II ability, on the other hand, is characterised by transformation and manipulation of the stimulus prior to making the response " (1970a, pp. 155-156).

In moving from a continuum of transformation to the definition of two levels of abilities, Jensen adopts a rather questionable classification of psychological processes. The problems inherent in this classification scheme are especially notable in the neglect of subject-initiated processing strategies. In regard to Level I abilities, use of this approach implies that one major group of mental abilities is characterised by little or no active structuring of information on the part of the individual.

RECENT RESEARCH ON MEMORY

If Level I ability is made up of registration, storage and retrieval capacities, then recent experimental research indicates that it does indeed involve some " elaboration, transformation, or manipulation of the input." Most current theoretical descriptions of the encoding, or registration processes of memory are in terms of ' elaboration,' ' rewriting,' ' transformation,' or ' attribute registration ' (Miller, 1956; Bower, 1972; Underwood, 1972). Even the more associative interpretations of encoding reject the idea that an encoding is simply a copy of an external event, and accept that it involves some manipulation of input (Voss, 1972, pp. 174-175). The studies by Wickens and colleagues (Wickens, 1972) suggests that while encoding may be carried out rapidly, it is both rich and complex in nature. Furthermore, the sequential interdependence of encoding, storage, and retrieval processes implies that any transformation of input during encoding influences the nature of what is stored and what is retrieved. Thus, recognition or recall ' with a high degree of fidelity ' does not necessarily imply an absence of transformation of input.

An interesting and deceptively simple task that can be viewed as a case in point, is the digit span subtest of the Wechsler intelligence scale. Jensen has used digit span as the purest test of Level I ability because, in his words, " reception and reproduction of the input with high fidelity is all that is required " (1970a, p. 156). While Jensen accepts that memory span is not a unitary ability, his use of digit span as his marker test for Level I implies that this test does not involve the subject in transformation of input.

Estes (1974) has noted that the processes involved in digit span performance should receive greater attention in future research because of the possible significance of these processes to theoretical conceptions of intelligence. This significance is reflected in research by Baumeister (1974) for instance, which indicates that the nature of processing in digit span is affected by conditions at the time of stimulus presentation and that input transformation, in the form of chunking or grouping, significantly influences mean levels of performance. Further, the rapid rate at which the digit span test is completed by subjects should not be used as an indicator of purely associative processes. As argued by Murdock (1974, p. 292), the fact that tasks such as digit span are carried out quickly does not necessarily mean that the processes involved are ' simple and direct '; the processes may be complex yet may operate efficiently in such tasks. In general, then, the evidence provided from experimental studies of serial learning suggests that apparently simple serial learning tasks often involve formation of sub-sequences which may be changed by subjects during completion of the tasks, and hence subject-initiated strategies are a significant variable in serial task performance (cf. Martin and Noreen, 1974).

This evidence may be complemented by studies of free recall. Jensen proposes that Level I and Level II ability can be distinguished with reference to free recall of uncategorised and categorised lists. Recall of uncategorised lists " requires nothing more than Level I ability, involving simply the reproduction of the input " (1970b, p. 91), while recall of categorised lists reflects Level II ability because it involves transformation in the form of clustering. This is a simplistic representation of a complex process which ignores evidence of clustering in both categorised and unrelated lists, and does not allow for the existence of organisational processes involved in free recall as evident, for example, in the studies of Pellegrino and Battig (1974). These studies indicated that both subjective and serial organisation increased over trials for both categorised and uncategorised lists.

FACTOR ANALYTIC STUDIES

In addition to the experimental evidence, correlational studies also have demonstrated the role of subject-initiated strategies in memory performance. Bergen *et al.* (1971) have demonstrated that different strategies can be detected by factor analysis, and that one source of these strategies is apparently the manner in which information is presented. Subjects in their study were found to adopt varying strategies of representing information in memory according to the complexity of the structure of the information displayed to them.

The detection of subject-initiated strategies in studies using factor analysis as a research tool raises interesting problems, which are highlighted in the case of Jensen's theory of Level I ability. If, as indicated experimentally, the individual's processing of stimulus information can be influenced by task demands, then the absolute use of one test as a reference point in factor analysis must be open to question. Similarly, if the standards of performance of subjects may be altered through the use of different processing strategies, as seems clear from rehearsal research (Brown, 1974), then the assumption that all subjects complete a test in the same way is also doubtful. The assumption of similar cognitive strategies in memory, and hence a de-emphasis on the strategies themselves, originates in factor analytic studies by the demonstration that memory tests form a separate factor. In the case of Jensen's theory, a factor analytic paradigm was used to verify the theory (Jensen, 1973), with three very similar memory tests used in the research.

But as noted by El-Koussy (see Guttman, 1967) correlations between tests can occur as a result of any combination of three characteristics of the tests:

> " Every test could be somehow thought of as having at least three main aspects—the content, the form, and the function. The content may mean numbers, words, pictures, figures, symbols, solids, actual things, situation or time. The form may mean the form in which the test items are present; i.e., classification, order opposites, analogies, . . . etc. The function may mean deduction, induction, memory, visualisation, manipulations . . . " (Guttman, 1967, p. 443).

While the examples of function given by El-Koussy may be questioned in the light of any particular theory adopted, the distinction of three aspects is useful. In Jensen's (1973) study, content and form were constant for the three tests, thus increasing the probability of high intercorrelations between the tests, and function was explicitly assumed to be the Level I processes as Jensen had defined them in earlier publications. When memory tests are grouped as noted here, and then included in a battery of other tests incorporating very different task demands, the memory tests quite naturally form a separate factor. It is easy to assume in this circumstance that strategies play no major role in task performance. It is only when memory abilities are explored within their own factor space, and variations are made in the content, form and function of the tests used, that subject strategies are clearly defined by factor analysis (e.g., Bergan *et al.*, 1971).

CONCLUSION

A curious lack of correspondence is evident between Jensen's theory of Level I ability and the current thrust of experimental and correlational research on memory. Indeed, the evidence noted here which does not match with Jensen's conception of Level I should be used in part to re-examine Jensen's complete theory of mental abilities. An alternative and more flexible approach to the nature of mental abilities should take into account both quantitative and qualitative differences in performance. Rather than distinguishing mental abilities solely on the basis of the presence of input transformation, it appears to be more useful, for theoretical and educational purposes, to consider the nature of the transforming processes. Such considerations appear to be explicit in the information processing approaches to cognitive abilities proposed by Carroll (1974) and Das, *et al.* (1975).

REFERENCES

BAUMEISTER, A. A. (1974). Serial memory span thresholds of normal and mentally retarded children. *J. educ. Psychol.*, 66, 889-894.

BERGAN, J. R., ZIMMERMAN, B. J., and FERG, M. (1971). Effects of variations in content and stimulus grouping on visual sequential memory. *J. educ. Psychol.*, 62, 400-404.

BOWER, G. H. (1972). Stimulus-sampling theory of encoding variability. In MELTON, A. W., and MARTIN, E. (Eds.), *Coding Processes in Human Memory.* Washington, D.C.: Winston.

BROWN, A. L. (1974). The role of strategic behavior in retardate memory. In ELLIS, N. R. (Ed.), *International Review of Research in Mental Retardation* (Vol. 7). New York: Academic Press.

CARROLL, J. B. (1974). *Psychometric Tests as Cognitive Tasks*: *A New Structure of Intellect.* Technical Report No. 4. Princeton, New Jersey: Educational Testing Service.

DAS, J. P., KIRBY, J., and JARMAN, R. F. (1975). Simultaneous and successive syntheses: An alternative model for cognitive abilities. *Psychol. Bull.*, 82, 87-103.

ESTES, W. K. (1974). Learning theory and intelligence. *Am. Psychol.*, 29, 740-749.

GUTTMAN, L. (1967). A psychological design for a theory of mental abilities. In JACKSON, D., and MESSICK, S. (Eds.), *Problems in Human Assessment.* New York: McGraw-Hill.

JENSEN, A. R. (1968). Patterns of mental ability and socioeconomic status. *Proc. Nat. Acad. Sci., USA*, 60, 1330-1337.
JENSEN, A. R. (1969). How much can we boost IQ and scholastic achievement? *Harv. Educ. Rev.*, 39, 1-123.
JENSEN, A. R. (1970a). Hierarchical theories of mental ability. In DOCKRELL, W. B. (Ed.), *On Intelligence*. Toronto: Ontario Institute for Studies in Education.
JENSEN, A. R. (1970b). A theory of primary and secondary familial mental retardation. In ELLIS, N. R. (Ed.), *International Review of Research on Mental Retardation* (Vol. 4). New York: Academic Press.
JENSEN, A. R. (1973). Level I and Level II abilities in three ethnic groups. *Am. Educ. Res. J.*, 10, 263-276.
MARTIN, E., and NOREEN, D. L. (1974). Serial learning: identification of subjective subsequences. *Cogn. Psychol.*, 6, 421-435.
MILLER, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.*, 53, 81-97.
MURDOCK, B. B. (1974). *Human Memory: Theory and Data*. Potomac, Maryland: Lawrence Erlbaum.
PELLEGRINO, J. W., and BATTIG, W. F. (1974). Relationships among higher order organizational measures and free recall. *J. Exp. Psychol.*, 102, 463-472.
UNDERWOOD, B. J. (1972). Are we overloading memory? In MELTON, A. W., and MARTIN, E. (Eds.), *Coding Processes in Human Memory*. Washington, D.C.: Winston.
VOSS, J. F. (1972). On relationship of associative and organizational processes. In TULVING, E., and DONALDSON, W. (Eds.), *Organisation of Memory*. New York : Academic Press.
WICKENS, D. D. (1972). Characteristics of word encoding. In MELTON, A. W., and MARTIN, E. (Eds.), *Coding Processes in Human Memory*. Washington, D.C.: Winston.

(*Manuscript received* 27th *April*, 1976)

CONFERENCE ABSTRACTS

Br. J. educ. Psychol., 47, 95-104, 1977

Papers presented to the Education Section of the British Psychological Society at York, 1976.

BURKE, ELIZABETH, M.Ed., Ph.D., A.B.Ps.S. Dept. of Education, University of Manchester.

CHILDREN'S STRATEGIES FOR RECOGNISING UNFAMILIAR WORDS

A study was reported of the strategies, or recognition cues, used by children at different stages of learning to read. Previous research (Smith, 1971; Goodman, 1965, 1972) has suggested that the immature reader is cued, or miscued, as he interacts with the written language, the major source of these cues being graphic symbols, syntax and semantics. The differential utilisation of these cues was investigated by an analysis of children's oral reading errors.

7, 8 and 9-year-old children were randomly selected from 12 county primary schools and were asked to read a story from Neale's Analysis of Reading Ability (Neale, 1958). Their errors, or miscues, were charted and scored according to their graphic proximity, and syntactic and semantic acceptability. Miscue score, as a measure of reading ability, was found to increase with age, but not uniformly for the categories of miscues, the age x miscue interaction being statistically significant. Scores in the graphic category showed a U-shaped trend, while the increase with age of score in the semantic category was particularly pronounced. Generally, pupils from informal schools were found to have the highest syntactic and semantic scores although the school type x miscue interaction was not statistically significant.

It is suggested that teachers be more sensitive to the kinds of miscues their pupils make, and that these miscues be treated as diagnostic aids in revealing the thinking strategies employed.

REFERENCES

GOODMAN, K. S. (1965). A linguistic study of cues and miscues in reading. *Elementary English*, 42, 639-643.

GOODMAN, K. S. (1972). Reading: the key is in children's language. *The Reading Teacher*, 25, 506-509.

NEALE, M. D. (1958). *Neale Analysis of Reading Ability*. London: Macmillan.

SMITH, F. (1971). *Understanding Reading: A Psycholinguistic Analysis of Reading and Learning to Read.* London: Holt, Rinehart and Winston.

BURLAND, J. S., B.Sc., Dip.Ed.Psych. Chelfham Mill School, Barnstaple, North Devon.

TEACHING APPROPRIATE BEHAVIOURS

The emphasis with many teachers is to simply extinguish inappropriate behaviour without regard to whether the child has the requisite skills to behave in an alternative manner. This paper essentially described ways of teaching these skills and how skill deficiencies should be assessed initially.

Alternative behaviours to temper tantrums, disrupting peers, not following instructions, bullying, answering back, hyperactivity and timidity were all discussed. Also discussed were social skills which improve a child's management of his environment, for instance, skills needed for effective social approach to adults, denial acceptance, interaction with the opposite sex, self-reinforcement and regulation.

EVANS, MARY. Director, Schools Council Project on the Education of Disturbed Pupils, Phillippa Fawcett College, London.

DEVIANT BEHAVIOUR IN THE CLASSROOM

A brief survey of the range of deviant behaviours encountered in the classroom was followed by an analysis of the teacher's role in terms of: (*a*) the qualities needed in a teacher; (*b*) the teacher's needs; (*c*) the role of the teacher *vis-a-vis* the role of the consultant. Consideration was given to those needs which can be appropriately met in the classroom. In conclusion, a query was raised as to whether there was a need to re-appraise philosophies about, and provision for, children showing deviant behaviour.

REFERENCES

CLARKE, A. M., and CLARKE, A. D. B. (1976). *Early Experience: Myth and Reality.* London: Methuen.

MORGAN, P. (1975). *Child Care—Sense and Fable.* Harmondsworth: Penguin.

FRANCIS, HAZEL, M.A., Ph.D. Dept. of Education, University of Leeds.

THE FIRST READER: LEARNING STRATEGIES OF 5-YEAR-OLDS

Few studies have been reported of the actual reading attempts of beginner readers. MacKinnon (1959), using the Richards-Gibson (1957) reading materials, was able to show the kinds of error made by 5-year-olds learning to grasp new words and read novel sentences. The materials were very carefully structured to maximise linguistic cues for successful attempts. Nevertheless, the errors were such that they might also be expected in other reading materials, but the syntactic/semantic context of words might well influence the distribution and frequency of such errors. A study by Soderbergh (1971) of the reading strategies of a pre-school child, also with specially designed materials, traced the development of grapheme pattern identification as a means of tackling new words.

This study reported the reading attempts made by 24 beginner readers from each of two schools representing widely different social backgrounds but using the same first 'reader.' Materials taken from this reader were used to vary the contexts in which errors occurred. The frequencies and types of error in different contexts were analysed with guidance from the MacKinnon and Soderbergh studies. The results were discussed with reference to strategies of reading through both visual patterning and linguistic structure.

REFERENCES

GIBSON, C. M., and RICHARDS, I. A. (1957). *First Steps in Reading English.* New York: Pocket Books Inc.

MACKINNON, A. R. (1959). *How Do Children Learn to Read*? Vancouver: Copp Clark.

SODERBERGH, R. (1971). *Reading in Early Childhood.* Stockholm: Almqvist and Wiksell.

HARRISON, COLIN, B.A., M.Phil. School of Education, University of Nottingham.

MEASURING READABILITY AT SECONDARY LEVEL: ASPECTS OF TEACHER JUDGMENT AND PUPIL RESPONSE

A set of 40 prose samples (24 first year, 16 fourth year) taken from a survey of texts in use in one school district during one week were assessed by 24 experienced teachers in terms of difficulty and interest level. Similar judgments were sought from first year (N=85) and fourth year (N=90) pupils in a different district. Finally, 200 first and fourth year children in a third group of schools were given 'cloze' tests of reading comprehension on the passages. The results of teacher and pupil judgments and mean cloze scores were compared with predictions of difficulty obtained from eight readability formulae.

Results overall indicated high correlations between readability formulae and pooled teacher judgments of difficulty (r=0·68—0·74 for all except one formula) and between different formulae (r=0·59—0·99). Teachers' judgments of interest level were significantly lower for fourth-year than for first-year passages, and correlated −0·49 with their estimate of passage difficulty. Teachers' and pupils' assessments of difficulty and interest levels were highly correlated (r=0·78—0·86). Correlations of cloze test scores with subjective criteria (r=0·10—0·69) were found to be lower than between those criteria and readability formulas.

JORDAN, R. R., B.Sc., M.Sc. Thomas Coran Research Unit, University of London Institute of Education.

CURRICULUM DEVELOPMENT FOR THE MENTALLY HANDICAPPED: CRITERIA FOR GOAL SELECTION

Curriculum planning involves specifying what to teach in both the long and the short term, but it must also consider methods and apparatus to be used and the context in which the teaching is to occur.

Observation of some of the differences between handicapped and non-handicapped development suggests a more structured teaching of some behaviours for handicapped children and the development of behaviour modification techniques has provided a powerful tool for this purpose. Such techniques have been used traditionally to train adaptive responses where there is a broad consensus on the relevance of the skill to the life of the individual.

But, just as at a broad level, we have moved from ideas of 'occupying' the mentally handicapped through 'training' and finally to 'education,' so there is a move towards teaching more general skills and especially those that will allow the handicapped individual to exert more choice in his own goal selection and that aim to give him the strategies for learning that are characteristically used by the non-handicapped.

Even for very specific aims such as feeding or receptive vocabulary, however, the choice of goals and the order in which they should be taught cannot be completely prescribed from knowledge of normal developmental sequences or from functional task analysis. Curriculum planning becomes experimental for each individual, taking his particular needs and the needs of his educators into account. When we come to areas of cognitive and social development, the sequency chosen will also reflect the theoretical framework we adopt and the choice of long-term goals in these areas are far more open to debate.

These difficulties, and the approaches adopted under their influence, were illustrated by some of the work over the last three years with Dr. C. C. Kiernan on developing teaching programmes and in working with parents and teachers on a variety of programmes with handicapped children.

REFERENCES

BRICKER, W. A., and BRICKER, D. D. (1973). Behaviour Modification Programmes. In P. MITTLER (Ed.), *Assessment for Learning in the Mentally Handicapped.* IR.M.M.H. : Churchill Livingstone.

KIERNAN, C. C. (1974). Application of Behaviour Modification in The Ward Situation. In A. C. GUNZBEY (Ed.), *Experiments in The Rehabilitation of the Mentally Handicapped.* IR.M.M.H. Butterworths.

KIERNAN, C. C. (1975). Experimental Investigation of the Curriculum for the Profoundly Mentally Retarded, for J. TIZARD (Ed.), *Mental Retardation: Concepts of Education and Research.* IR.M.M.H. Butterworths.

MARTIN, BARRY, B.A., B.Ed., and WHELDALL, KEVIN, B.A., A.B.Ps.S. School of Education, University of Birmingham.

SOCIO-ENVIRONMENTAL INFLUENCES ON THE RECEPTIVE LANGUAGE DEVELOPMENT OF 'RISING FIVES'

This report is concerned with the various influences, which may broadly be termed 'socio-environmental,' on the development of receptive language in young children starting school. The variables investigated during the study included social class, nursery education, and various other family variables and their effects on the English Picture Vocabulary Test and the Sentence Comprehension Test, which is currently being developed (Wheldall, 1976).

In a previous paper to the Education Section (Wheldall, 1974) we presented evidence to show that by no means all children starting school could demonstrate comprehension of 'well formed' sentences as the psycholinguists had suggested, and that nursery education was shown to exert a powerful beneficial effect on the receptive language development of working class children. However, a follow-up study two years later confirmed the well documented 'washing out' effect; nursery educated children were no longer in advance of their non-nursery educated peers on measures of receptive language development.

These findings are extended and reconsidered in the light of the results from our most recent study which consists of a large scale investigation of the various influences on the receptive language performance of 100 5-year-olds starting school. Each child's mother was interviewed concerning various aspects of family environment and the child's pre-school educational experience. Social class differences were confirmed but the interaction between class and pre-school provision proved to be more complicated than we had previously hypothesised. Working class children appeared to benefit more (in terms of improved receptive language) from pre-school provision than either middle class or lower working class children but our previous study casts doubt on the long-term beneficial effects.

REFERENCES

WHELDALL, K. (1974). Social factors affecting the comprehension of pre-school children. Paper presented to the Annual Conference of the Education Section of the British Psychological Society, Edinburgh, September, 1974.

WHELDALL, K. (1976). Receptive language development in the mentally handicapped. In BERRY, P. (Ed.), *Language and Communication in the Mentally Handicapped.* London: Arnold.

MILES, ROGER, B.Sc., M.Phil., A.B.Ps.S. N.D.P.C.A.L., London.

NATIONAL DEVELOPMENTS IN COMPUTER ASSISTED LEARNING

The National Development Programme in Computer Assisted Learning (NDPCAL), which began its five-year life on 1st January, 1973, with a budget of up to £2 million, is an unusually large innovation in education and has a number of novel features. For example it is directly controlled by government and has developmental not research aims.

The Programme has funded some 30 projects which are using the computer in:

(*a*) tertiary education, in mathematics and statistics, science, engineering, medicine and teacher education;

(*b*) secondary education, in mathematics, biology, geography, history and school timetabling;

(*c*) primary education, a computer managed remedial reading scheme;

(*b*) military and industrial training.

This paper gave an overview of the educational projects concentrating on those likely to be of most interest to educational psychologists. The varied roles which the computer can take in support of the educational process were described and a number of problems and issues discussed.

Computer assisted learning (CAL) is growing vigorously in secondary and tertiary education and so there is a good chance that educational psychologists will meet it during their work. There is scope for psychologists to help with the development and evaluation of CAL activities or pick up some of the research questions which have arisen.

REFERENCES

HOOPER, R., and TOYE, I. (Eds.) (1975). *Computer Assisted Learning in the United Kingdom: Some Case Studies.* London: Councils and Education Press.

MITTLER, PETER, M.A., Ph.D., F.B.Ps.S. Hester Adrian Research Centre, University of Manchester.

PSYCHOLOGY AND SPECIAL EDUCATION

Since special education is currently under scrutiny by the Warnock Committee, it seems appropriate to raise a number of questions about the contribution which psychology might make to the development of educationally handicapped children.

The BPS evidence to Warnock was briefly summarised in the *Bulletin* (January, 1976), but had for reasons of space to omit an appendix which attempted to summarise some aspects of the relationship between psychology and special education. The Introduction to this appendix is reproduced here in full.

INTRODUCTION

" 1.1 The primary role of the psychologist within special education is to apply the knowledge accumulated in the fields of general, experimental and developmental psychology to the problems presented by the child, his family or the school. Psychologists are trained in child development, experimental methodology, psychological assessment and psychological methods of intervention, as well as in the more general areas of neuro-psychology, and the biological basis of development, both normal and abnormal. From a synthesis of training in these areas, psychologists bring a unique psychological perspective to bear on the problems presented by developing children. This perspective emphasises the *interaction* between the child and his physical and social environments. Thus, psychologists seek for understanding and explanations of presenting problems in the interaction between the situation in which the child finds himself and what the child brings with him to that situation. Solutions to problems are, therefore, looked for both in altering the environment and on working more directly with the individual child."

The implications of this statement were elaborated.

MULLER, D. J. F., B.Ed. School of Speech Pathology, Leicester Polytechnic.

EXPERIMENTAL INVESTIGATIONS OF CHILDREN'S CONCEPTS OF PROPORTION

Bryant (1974) has suggested weaknesses in Piaget's analysis of cognitive development (see Piaget and Inhelder, 1966). Furthermore, investigations of the concept of proportion have produced conflicting data. Bryant has suggested that 4-year-olds understand proportion whereas Piaget has argued this is not true until the stage of formal operations (11+).

Two experiments were presented to resolve this conflict and to answer two basic questions. Firstly, depending on the extent to which children can analyse an initial operation can they make inferences to link two such operations? Secondly, are different levels of understanding the concept of proportion identifiable based on Bryant's thesis that relative codes develop before absolute codes, that is quantification occurs later in development.

Experiment I refined Bryant's initial work answering the first question positively. Experiment II showed that different levels supporting Bryant's thesis could be identified. However, a higher level of functioning was proposed, extending Bryant's thesis to include a combination of relative and absolute coding.

These experiments illustrated the validity of Bryant's criticism of Piaget's work and extended this thesis. Children can make inferences and link perceptual experiences and acquire the ability to register relative values before quantifying them in absolute terms.

REFERENCES

BRYANT, P. (1974). *Perception and Understanding in Young Children.* London: Methuen.

PIAGET, J., and INHELDER, B. (1966). *The Psychology of the Child.* Routledge and Kegan Paul.

NELSON, DIANA M. Meldreth Manor School, Meldreth, Herts.

CHANGES IN THE BEHAVIOUR OF A TEACHER DEALING WITH AGGRESSIVE CHILDREN WHEN GIVEN A RESTRUCTURED TEACHING ENVIRONMENT AND GUIDANCE IN OPERANT TECHNIQUES

This presentation was in two parts. A 15-minute videotape demonstrated one subject's behaviour, the strategies adopted to initially contain the behaviour and the training used to modify it.

Having set the scene an analysis of changes in the teacher's and the child's behaviour followed.

(*a*) A consideration of the problems posed by aggressive children—with special reference to the subject was made.

(*b*) A brief description of the two types of tantrums exhibited by the subject was undertaken.

(*c*) The structural changes which were the key to the treatment were summarised.

(*d*) A detailed analysis of the shift in the teacher's approach from one of ' appeasement ' to that of training the subject to work productively in the classroom was outlined.

(*e*) An analysis of the changes in the subject's in-class performance was presented.

NEVILLE, MARY, H., Ph.D. School of Education, University of Leeds.

THE TRANSITION TO SILENT READING

For the English child in the first school, the initial mode of reading is oral. In the development of reading competence three stages, which overlap to some extent, have been suggested by Goodman and later designated by Pugh as oral, aural, and silent. The role of all stages is explained more fully elsewhere (Daly, *et al.*, 1975), but the aural stage of ' listening in ' to the text appears to be important in helping children develop from oral to silent readers. If they remain at the oral stage, they may read in a word-by-word sequential manner, neither making use of context effectively nor reading ' predictively.'

Two investigations (Neville and Pugh, 1974, and in press) with 9-10-year-old subjects performing cloze tests of silent reading, listening, and restricted reading (where the sequential presentation constraints of the listening test were similar in some respects) seem to indicate a different type of reading for better and poorer readers. Poorer readers appeared to relate their reading comprehension only to the word they had reached and to preceding context (as in the listening or restricted reading tasks). These children seem to gain some benefit from using reading while listening materials (Neville and Pugh, 1975) may assist them to move to the aural stage and help them break ' visual-to-name ' translation habits of reading as they begin to adopt strategies more appropriate to true silent reading.

REFERENCES

DALY, B. V., NEVILLE, M. H., and PUGH, A. K. (1975). *Reading while Listening: An Annotated Bibliography of Materials and Research*. Leeds: University of Leeds Institute of Education.

NEVILLE, M. H., and PUGH, A. K. (1974). Context in reading and listening: a comparison of children's errors in cloze tests. *Br. J. educ. Psychol.*, 44, 224-232.

NEVILLE, M. H., and PUGH, A. K. (1975). An empirical study of the reading while listening method. In D. MOYLE (Ed.), *Reading: What of the future*? London: Ward Lock.

NEVILLE, M. H., and PUGH, A. K. (in press). Context in reading and listening: variations in approach to cloze tasks. *Reading Res. Q.* (1) 12.

PUGH, A. K., B. A., M.Phil. Faculty of Educational Studies, The Open University.

SILENT READING STRATEGIES

True silent reading differs from earlier approaches to reading in that it involves the use of a wider and different range of techniques and types of strategy for achieving purposes. These techniques are impossible in other modes of reading (Pugh, 1975). It is unfortunate that speed of reading has been the main concern of most teachers who have concerned themselves with silent reading, for speed of reading is very difficult to measure and there are grounds for thinking that many apparent gains in reading speed after speed-reading courses are a result of learning how to do the test exercises used. A more important weakness is that the courses and some of the literature ignore the techniques and strategy appropriate to effective adult silent reading, because their model of reading is what has been designated elsewhere as aural reading.

The purpose of this paper was to review a number of studies in which the writer has attempted to delineate and assess silent reading techniques and strategies. For this purpose apparatus was devised for monitoring subjects' use of books for set purposes. Subjects were middle school children (Neville and Pugh, 1975, and forthcoming), sixth-formers (Pugh, 1976) and university students. Scores from standardised tests of reading were available for all subjects, and results of tests of reading and measures of intelligence in some of the studies. The findings suggested:

(*a*) that certain techniques can be identified;

(*b*) that the ability to select the appropriate strategy (i.e., context-bound and goal-directed combination of techniques) is not closely related to scores on other measures;

(*c*) that subjects found the tasks more difficult than their teachers, explicitly or implicitly, seemed to expect;

(*d*) that how best to teach a strategic approach remains a problem for further research.

REFERENCES

NEVILLE, M. H., and PUGH, A. K. (1975). Reading ability and ability to use a book: a study of middle school children. *Reading*, 9, 23-31.

PUGH, A. K. (1975). The development of silent reading. In W. LATHAM (Ed.), *The Road to Effective Reading*. London: Ward Lock.

PUGH, A. K. (1976). Implications of problems of language testing for the validity of speed reading courses. *System*, 4, 23-31.

Shayer, M. Concepts in Secondary Mathematics and Science, Chelsea College.

DISTRIBUTION OF PIAGETIAN STAGES IN CHILDREN 9–16

Piaget's description of the development of thinking in adolescence is very attractive to someone interested in curriculum design in science (Shayer, 1970) because it can be used to measure both the level of development in children and the level of conceptual demand of science activities. This paper described work on the former of these two measurements. Clearly, in an ideal curriculum, these levels should match each other, but to plan this on behalf of pupils one needs to know how the levels of thinking which Piaget describes are distributed in the school population.

The proportion of children showing early and late concrete operational thinking, and early and late formal operational thinking, was measured in a sample of 10,000 children between the ages of 9 and 14. The test instruments were a form of group test called Class Tasks, derived from the individual interview situations described by Piaget. These tests indicate that most children in early adolescence showed rapid development of concrete operational thinking, but that only one-fifth of the children showed the further development of formal operational thought. The representativeness of the findings was ensured by relating the distribution of Piagetian stages at each age-level to the norms of a standardised non-verbal reasoning test (Shayer *et al.*, 1976). Susequently, this measure has been extended to include fourth- and fifth-year pupils.

These results were then used to examine the suitability of the Nuffield Combined Science course for a school population below that of the top 20 per cent.

REFERENCES

Shayer, M. (1970). How to assess science courses. *Educ. Chem.* 7 (5) 182.

Shayer, M., Küchemann, D. E., and Wylam, H. (1976). The distribution of Piagetian stages in British middle and secondary school children. *Br. J. educ. Psychol.*, 46, 164-173.

Wheldall, Kevin, B.A., A.B.Ps.S., and Poborca, Barbara, J., B.A., D.C.P. School of Education, University of Birmingham.

CONSERVATION WITHOUT CONVERSATION ? TOWARDS A NON-VERBAL PARADIGM

There has been considerable interest and criticism surrounding the nature and effects of the various spoken instructions on Piagetian conservation tasks, which Peill (1975) has recently reviewed. It appears that some of the problems might be solved if a non-verbal procedure could be devised which required neither comprehension of verbal instructions nor verbal justification of the child. Braine's (1962) early plea for such an approach has prompted many attempts, reviewed by Miller (1976).

The experimental work reported here constituted an attempt to devise a non-verbal paradigm for testing conservation of liquid, based upon an operant discrimination learning situation with similarities to a go/no go procedure. The child's task was to learn to press or to refrain from pressing a button contingent upon whether two glass jars contained similar or different amounts of water. Button-pressing to similar amounts and refraining from pressing to different amounts was reinforced by means of a ' smartie ' dispenser. By these means the child was taught to signal his evaluation of the relationship between the two amounts of water in response to a non-verbal request. When the child reached criterion, we transformed one of two quantities of water previously signalled as similar into, say, a taller, narrower jar and again non-verbally requested an evaluatory response to the transformed stimulus.

Our initial results appeared to demonstrate that young children who could not conserve with the usual verbal procedure were more likely to be able to demonstrate conservation within this non-verbal paradigm.

REFERENCES

BRAINE, M. D. (1962). Piaget on reasoning: a methodological critique and alternative proposals. *Monographs of the Society for Research in Child Development*, 27 (2, Serial No. 83).

MILLER, S. A. (1976). Non-verbal assessment of Piagetian concepts. *Psychol. Bull.*, 83, 405-430.

PEILL, E. J. (1975). *Invention and Discovery of Reality: The Acquisition of Conservation of Amount.* London: Wiley.

WILKINSON, J. E., B.Sc., M.Ed., A.B.Ps.S., and GRANT, SISTER DOREEN, M.Ed., C.B.E. Dept. of Education, University of Glasgow, *and* Convent of Notre Dame *respectively.*

THE STRATHCLYDE EDUCATIONAL ACTION RESEARCH PROJECT

Recent surveys carried out by the National Children's Bureau and the Office of Population Census and Surveys indicated the extent of social disadvantage in the West of Scotland—in particular Glasgow. The degree to which children living in disadvantaged communities underperform in school was reported in *From Birth to Seven.* To date, expensive projects such as Headstart and EPA have achieved limited success in promoting the education of children living in difficult social conditions.

However, recent small-scale studies offer a more hopeful note (Donachy, 1976; Wilkinson and Murphy, 1976). A new consciousness by professionals of the power, competence and importance of parents is called for.

Working to expand parents' conceptualisation of their environment, a preliminary programme has been carried out with parents living in a small community in Glasgow. A joint Strathclyde Region/Glasgow University Project has been established incorporating this initial work with funds from the Region, the Urban Renewal Unit and the Notre Dame Trust. The emphasis will be on an intensive programme of short-term help to assist a community to take more responsibility for its progress.

REFERENCES

DONACHY, W. (1976). Parent participation in pre-school. *Br. J. educ. Psychol.*, 46, 31-39.

WILKINSON, J. E., and MURPHY, H. F. (1976). Differential methods enhancing cognitive growth in inner-city pre-school children. *Child Care: Health Dev.*, 1, 1-11.

VAN DER WILL, CHRISTINE, B.A., B.Phil. School of Education, University of Birmingham.

THE WORDING OF SPOKEN INSTRUCTIONS TO CHILDREN AND ITS EFFECTS ON THEIR PERFORMANCE OF TASKS

Little experimental attention has been given to the linguistic formulation of spoken instructions to children in a task-oriented situation, and to the measurement of their effectiveness. Amidon and Carey's (1972) study made an exception here, demonstrating that instructions comprised very powerful independent variables. Abercrombie (1960) also stressed the importance of instructions in a teaching situation.

This investigation aimed to measure children's understanding of spoken instructions, in which sentence length and complexity were systematically varied, by means of giving and scoring some simple drawing tasks, which do not involve decisive variables such as mathematical or writing ability; their outcome can, therefore, be seen to be due almost unequivocally to the language of the instructions themselves.

Two experiments were reported, both involving 8-year-old children from normal junior schools. In the first experiment, 36 children were tested for latency and accuracy of response to the differently worded instructions. Their responses to the longer instructions were significantly faster, but not significantly more accurate than to the shorter instructions. The second experiment, involving 30 children, tested whether a slower presentation of the shorter instruction would make it as effective as the longer one. This time, whilst there was no significant effect on speed, children's accuracy was significantly better for the longer instruction.

REFERENCES

ABERCROMBIE, M. L. J. (1960). *The Anatomy of Judgment*. Harmondsworth: Penguin.
AMIDON, A., and CAREY, P. (1972). Why 5-year-olds cannot understand before and after. *J. verb. Learn. verb. Behav.*, 11, 417-423.

CRITICAL REVIEW

Br. J. educ. Psychol., 47, 105-107, 1977

HOW MANY TEACHING STYLES ? HOW MUCH PROGRESS ?

BY J. A. WILSON
(NICER Research Unit, Queen's University, Belfast)

BENNETT, S. N. (with JORDAN, J., LONG, G., and WADE, B.) (1976), *Teaching Styles and Pupil Progress.* London: Open Books, pp. xiii+201, p. £2·95, c. £5·75.

Teaching Styles and Pupil Progress, a research report by Bennett and his co-workers from the Department of Educational Research at Lancaster, describes how a survey of teaching practices formed the basis of a study of the effect on pupils' cognitive achievement of formal, informal and mixed teaching styles. In his foreword to the study Entwistle acknowledges the controversial nature of the enquiry, and, in view of its sensitive relationship to the progressive-traditional debate, is at pains to declare the interests of the research team. None of its members, we are assured, came to the project as a dyed-in-the-wool traditionalist.

In the survey a questionnaire on six areas of classroom teaching practice was completed by 1,258 teachers of third- and fourth-year pupils in a sample of some 800 junior primary schools in the North-West of England. The 19 items which most clearly defined seven factors in the replies of 468 fourth-year teachers were used to cluster these teachers by reported teaching style. The teacher clusters which emerged from the cluster analysis were then subjectively ordered on a formal-informal continuum.

A sub-sample of 37 teachers was chosen on the basis that their responses closely matched the central profiles of seven of the 12 clusters. Twelve of the 37 teachers were selected from the two clusters at the formal extreme of the continuum, and 13 were selected from the two clusters at the informal extreme. The remaining 12 teachers were drawn from three clusters chosen to represent intermediate or ' mixed ' teaching styles.

These three groups of teachers, representing formal, mixed and informal teaching styles, constituted the levels of treatment in a quasi-experimental analysis of pupil performance in reading, English and mathematics, as measured by end of fourth-year objective test scores regressed on scores for the same tests taken one year previously. Analysis of covariance for the 950 pupils in the 37 classrooms showed that pupils taught by formal and mixed styles showed significantly superior progress in reading as compared with those pupils taught informally. In mathematics and English the formally taught pupils were superior to mixed and informally taught pupils. On all criteria boys, but not girls, with low initial achievement under-achieved in formally-taught classrooms, whereas boys and girls with high initial achievement over-achieved in formally taught classrooms. Since most pupils, irrespective of personality profile, showed greater progress in formal classrooms, Bennett concluded that teaching style had a more powerful effect on pupil progress than had pupil personality.

A further analysis of samples of imaginative and descriptive writing indicated that pupils in formal and mixed classes were better at punctuation than, but not significantly inferior in creative and imaginative writing to, pupils in informal classes. Structured observation of a small sample of pupils in formal and informal classrooms showed that pupils engaged in more work activity in formal classrooms, whereas pupils in informal classrooms interacted with each other to a greater extent in the context of both work and social activities.

The report, which has received considerable publicity in the press and on radio and television, is clearly aimed at a wide readership. The result is a readable account which is lacking in technical information. Time and again findings are alluded to rather than evidenced. We are told, for example, that the questionnaire responses and typologies of the third-year and the third- and fourth-year combined teachers were ' strikingly similar ' to those of the fourth-year teachers whose profiles are described here and elsewhere (Bennett and Jordan, 1975). We are told that ratings based on the questionnaire items and cluster descriptions, and made available independently by research staff and primary school advisors, agreed closely with the cluster descriptions of the 37 teachers whose classes were involved in the experiment. We are not given the opportunity to examine these consistencies for ourselves.

Bennett's teachers were classified by teaching style on the basis of their responses to a questionnaire. This procedure has been criticised by Powell (1976) on the grounds of its failure to sample teaching through skilled and systematic observation of classroom practice. Certainly, it must be emphasised that Bennett's analysis of teaching is based on what teachers *professed* to practise, rather than on what they were *observed* to practise, though attempts were made to substantiate the former observationally. At the same time it may be argued that until such time as the techniques of systematic observation succeed in pin-pointing those classroom processes which most effectively promote cognitive growth on the part of pupils we must continue to pay attention to teachers' reports of how they organise classroom learning. To establish meaningful links between reported practice and pupil growth is valuable in its own right. That it may also serve to indicate areas of classroom practice for further intensive observation is not, however, a central issue in Bennett's study.

In the introductory sections of the report Bennett is at pains to question the value of a dichotomous or uni-dimensional view of teaching. Despite this, and despite also the evidence from the factor analysis, which was essentially multi-dimensional, Bennett proceeded to place his 12 teaching profiles in an order of formality, preparatory to collapsing seven categories into the three broad types of formal, mixed and informal. Since no proof, such as that of hierarchical clustering, is presented which would support the ordering and collapsing of clusters in this way, one is forced to conclude that this treatment of the data runs counter to both previous argument and evidence.

Indeed, Bennett himself comes close to admitting to similar reservations when he refers (p. 157) to an analysis based on the seven styles from which his three broad teaching styles were derived. Apparently this analysis showed that three of the seven styles were associated with progress above that expected from initial achievement These were the two ' formal ' styles and one ' mixed ' style which, in Bennett's words, " conjoined a form of integrated day with strict teacher control of content and pupils." This return to an examination of the more clearly differentiated clusters, in conjunction with an inspection of the profile of his successful informal teacher, leads Bennett to suggest that " careful and clear structuring of activities together with a curriculum which emphasises cognitive content are the keys to enhanced academic progress " (p. 160). His further conclusion that curriculum organisation is more important for pupil performance than classroom organisation may be more revealing about teaching processes than any conclusion about the comparative efficiency of formal and informal teaching styles.

In an *ex post facto* design the experimenter, having no control over the assigning of pupils or classrooms to treatments, should statistically adjust the dependent variable for any extraneous condition suspected of interacting with a treatment. Bennett rightly uses the pre-test measures as covariates. We are also told, in the second edition of the report, that there were no significant differences between the sample of formal and informal teachers on age, sex, experience, class size and age of school, and that homogeneity of regression was established for the three experimental groups prior to the analysis of covariance. Despite these reassurances he has neglected to adjust for such an obviously correlated side-effect as the 11-plus examination, which was least prevalent in the informal schools. Also, given that the formally taught pupils had the highest mean scores on the pre-tests, he might usefully have looked for other context variables, such as father's occupation, which even at a gross level may well have been related by way of parental choice to a ' formal=successful=reputable ' set of predispositions. There is, of course, a limit to the extent to which covariance adjustments can compensate for lack of experimental control. That limit is inadequately exploited in the present study.

typologies of the third-year and the third- and fourth-year combined teachers were ' strikingly similar ' to those of the fourth-year teachers whose profiles are described here and elsewhere (Bennett and Jordan, 1975). We are told that ratings based on the questionnaire items and cluster descriptions, and made available independently by research staff and primary school advisors, agreed closely with the cluster descriptions of the 37 teachers whose classes were involved in the experiment. We are not given the opportunity to examine these consistencies for ourselves.

Bennett's teachers were classified by teaching style on the basis of their responses to a questionnaire. This procedure has been criticised by Powell (1976) on the grounds of its failure to sample teaching through skilled and systematic observation of classroom practice. Certainly, it must be emphasised that Bennett's analysis of teaching is based on what teachers *professed* to practise, rather than on what they were *observed* to practise, though attempts were made to substantiate the former observationally. At the same time it may be argued that until such time as the techniques of systematic observation succeed in pin-pointing those classroom processes which most effectively promote cognitive growth on the part of pupils we must continue to pay attention to teachers' reports of how they organise classroom learning. To establish meaningful links between reported practice and pupil growth is valuable in its own right. That it may also serve to indicate areas of classroom practice for further intensive observation is not, however, a central issue in Bennett's study.

In the introductory sections of the report Bennett is at pains to question the value of a dichotomous or uni-dimensional view of teaching. Despite this, and despite also the evidence from the factor analysis, which was essentially multi-dimensional, Bennett proceeded to place his 12 teaching profiles in an order of formality, preparatory to collapsing seven categories into the three broad types of formal, mixed and informal. Since no proof, such as that of hierarchical clustering, is presented which would support the ordering and collapsing of clusters in this way, one is forced to conclude that this treatment of the data runs counter to both previous argument and evidence.

Indeed, Bennett himself comes close to admitting to similar reservations when he refers (p. 157) to an analysis based on the seven styles from which his three broad teaching styles were derived. Apparently this analysis showed that three of the seven styles were associated with progress above that expected from initial achievement These were the two ' formal ' styles and one ' mixed ' style which, in Bennett's words, " conjoined a form of integrated day with strict teacher control of content and pupils." This return to an examination of the more clearly differentiated clusters, in conjunction with an inspection of the profile of his successful informal teacher, leads Bennett to suggest that " careful and clear structuring of activities together with a curriculum which emphasises cognitive content are the keys to enhanced academic progress " (p. 160). His further conclusion that curriculum organisation is more important for pupil performance than classroom organisation may be more revealing about teaching processes than any conclusion about the comparative efficiency of formal and informal teaching styles.

In an *ex post facto* design the experimenter, having no control over the assigning of pupils or classrooms to treatments, should statistically adjust the dependent variable for any extraneous condition suspected of interacting with a treatment. Bennett rightly uses the pre-test measures as covariates. We are also told, in the second edition of the report, that there were no significant differences between the sample of formal and informal teachers on age, sex, experience, class size and age of school, and that homogeneity of regression was established for the three experimental groups prior to the analysis of covariance. Despite these reassurances he has neglected to adjust for such an obviously correlated side-effect as the 11-plus examination, which was least prevalent in the informal schools. Also, given that the formally taught pupils had the highest mean scores on the pre-tests, he might usefully have looked for other context variables, such as father's occupation, which even at a gross level may well have been related by way of parental choice to a ' formal=successful=reputable ' set of predispositions. There is, of course, a limit to the extent to which covariance adjustments can compensate for lack of experimental control. That limit is inadequately exploited in the present study.

BOOK REVIEWS

FALMAGNE, R. J. (Ed.) (1975). *Reasoning: Representation and Process.* Hillsdale, New Jersey: Lawrence Erlbaum Associates, pp. xii × 275, £10·15

To readers inured to a steady stream of edited productions composed of papers having little in common but shared hard covers this volume will be a pleasant surprise. It had its origin in a symposium in 1973 when the editor became convinced that if integration of the empirical findings and the theoretical positions emerging from research on logical thinking was to be achieved it was necessary to pull together the contributions scattered in the literature in a unified presentation. The main aim of the book is to provide a unified, if necessarily transient, picture of the current state of research on the main sub-areas of logical thinking—propositional reasoning, syllogistic reasoning and transitive inference.

In an introduction the editor distinguishes two traditions in the study of logical reasoning. The ' propositional ' tradition focuses on verbally stated, self-contained problems such as syllogisms and is mainly concerned with adult subjects while the Piagetian tradition involves asking children to reason about essentially concrete material. Despite these differences the focus of theoretical interest in both cases is to characterise the processes which lead the child or adult to behave competently in the situations considered. The same range of viewpoints on this issue can be discerned in both traditions. The subject's behaviour can be regarded as attributable to the operation of generally applicable logical processes or, in contrast, to the application of pragmatic content bound rules perhaps involving the use of representational techniques such as imagery to provide a concrete model of the problem. The papers included in the volume illustrate the relevance of this issue in studies of adult reasoning, ' propositional ' reasoning in children and children's reasoning in Piagetian situations.

In the first of four papers concerned with adult reasoning, Johnson-Laird presents an information-processing model covering lexical and propositional reasoning and reasoning with quantifiers. He concludes that if discourse is supplemented by knowledge made available through commonsense assumptions, it is unnecessary to detract from the generality of the model by postulating special rules of inference to deal with specific topics, although the content of statements may exert a selective bias on the availability of different rules of inference. Staudenmayer accepts that general, logical rules govern the process of making or evaluating inferences but on the basis of experimental study of the relative effect of abstract or concrete material on propositional reasoning argues that the interpretive process that precedes evaluation is influenced by specific knowledge and presuppositions about context. In similar vein, Revlis reports experimental results in support of the view that deductive errors on categorical syllogisms are only indirectly affected by the problem's content and that errors are attributable to the reasoner's misrepresentation of the problem's propositions rather than insufficiencies in logical skills. The nature of the representation of the initial information in logical tasks provides the theme for papers by Neimark and Scribner. The latter reports a cross-cultural study involving the comparison of the performance of illiterate West Africans on memory and inferential tasks involving syllogisms expressed in concrete terms. The results blur the distinction between the successive phases of encoding and inference operations which figure in temporal stage models of logical reasoning. The memory data suggest that encoding premises involves logical operations of inference just as the results of inference depend on the outcome of encoding.

Studies of ' propositional ' reasoning in children are represented by papers by Osherson and Falmagne. Osherson's contribution is a thumbnail sketch of his well-known model in which reasoning operates on the form of the statements regardless of their content. The three experiments reported by Falmagne arise from a process model of children's performance on propositional tasks and involve the testing of hypotheses regarding the relationship between evaluation strategies and forms of representation by means of latency and error data. Representation, also, provides the focus for Trabasso's discussion of the development of transitive inference in which he presents convincing experimental evidence that both children and adults use a common underlying representation—a linear order of a spatial nature—to integrate comparative relations among events. The final developmental paper by Youniss is, also, concerned with Piagetian situations and reports four experiments concerned with the sequence of development of seriation, transitivity and measurement of length.

In a final overview Falmagne considers the status of the conceptual framework underlying the study of logical reasoning in which the two functions of representation of information and logical processing are regarded as being distinct even if not identified as serially organised stages. She concludes that a more dialectic scheme is called for in which they are seen as alternative perspectives on the same object rather than complementary ways of slicing it. It seems probable that the increasing application of information processing analysis in both propositional and Piagetian studies of logical reasoning will produce such a revision of the conceptual framework since there is an increasing emphasis on a procedural definition of knowledge in information processing work. At any rate discussion of this prescription, as of much else in this volume, will provide stimulating activity for the postgraduate students who are its intended readership.

J. G. WALLACE.

MALMQUIST, E., and GRUNDIN, H. U. (1975). *Educational Research in Europe—Today and Tomorrow*. Lund, Sweden: C.W.K. Gleerup/Liber Laromedel, pp. 442, 50 sw. cr.

This is a massively dull book. I suppose I am still trying to recover from reading it, and suffering the pent-up exasperation of one who was foolish enough to expect an interesting, definitive study on an important subject. What, in fact, we have in this volume is a distilled-water account of educational research; the book never gets within sight of a youngster in a classroom or a truant enjoying his day off.

It really is all very sad. So much hard work has gone into it, there has been so much sifting of replies, that one feels bound to ask what has gone wrong. Let us see what it is trying to do.

The study is part of the project, " Plan Europe 2000," administered by the European Cultural Foundation and attempts to assess the status of European research in the field of education. The work begins with a review of 1,000 items from the relevant literature, of which 45 per cent is from America. Three questionnaires were then constructed and sent to researchers, research directors and research financing agencies in seven European countries. We have no replies from the Soviet Union so we are left with six.

The result, if not indeed the aim, is to produce endless correlational tables indicating how one country compares and differs from another in such matters as the importance of various fields of study for the training of educational researchers, whether there is too much or too little self-criticism in educational research, and so on. The reader is left with a mass of data and no clear picture of what is going on. The odd quotation such as, " In Italy research in education rests within the domain of spontaneous initiative " throws more light than pages of the statistics.

It appears then that the level of the book is not in keeping with its somewhat grandiose aims. It should not be necessary in a volume of this kind to spend many pages spelling out the definitions of terms such as organisation, the differences between pure and applied research, or even six pages quoting definitions of educational research. Again, if the aim is "to analyse the status of European research in the field of education and identify significant trends and problems as well as means and ways of coping with these problems " (p. 17), then it cannot be done by concentrating upon the kinds of question in this study such as the number of reports put out by each country in printed and mimeographed form. We can be expected to know that " the number of copies of each report varies" whilst the fact that Italy puts out twice as many mimeographed reports as other countries is hardly going to change the course of research history.

I would concur with a quotation from this book that " a body of literature can grow faster than a body of knowledge," and that educational research too often demonstrates this trend. The present book makes its own contribution. But it was of note that even when the questionnaire had an imaginative burst and, uncharacteristically, asked what possible future discoveries if made before the year 2,000 might be designated the educational discovery of the century, nearly half the respondees did not know. Inevitably some took the opportunity to be enigmatic, " N'est il pas plus prudent de se taire ? ", or funny, " I wish I knew the answer to this ! " But it does strike the writer as distinctly odd that in a field where the major issues are so challenging, when we are trying so hard to grapple with such worthwhile issues as the perceptual analysing mechanisms of the nervous system, the storage and retrieval characteristics of the brain, the basis of linguistic operations (to mention only a few), there should be no eagerness to discuss priorities. Educational research should be at a stage where it could make an enormous advance as genetics and biochemistry have already done.

As in the case of other sciences our research has gone through its long period of data collection and followed interests in a variety of subjects such as mental testing, moral development and delinquency, etc. But today, there are fundamental issues such as the development and operation of the sensory and storage mechanisms thereby enabling a biological system to respond and adapt to its environment, which could be tackled with hopes of success. A subject needs to come of age, to be aware what it might achieve and so state its priorities and get on with them.

HARRY KAY.

PARANJPE, A. C. (1976). *In Search of Identity.* New York: Wiley, pp. vii+263, £8·30, $15·40.

Dr. A. C. Paranjpe studied and taught at the University of Poona and subsequently worked in the College Youth Project in that city. He has also spent a year with Erik Erikson, and since 1976 has been at Simon Fraser University. This book arises from all these experiences.

The starting point is the belief that: " The natural progress of the lives of people is one of the highly neglected aspects of the study of man." There are some exceptions to this, as Paranjpe points out, for example the studies of Terman and of White. The present book is a contribution to this area, using a particular theoretical orientation.

The book begins with a rather useful survey of several different relevant approaches to the understanding of human beings, including the influence of developments in sociology and anthropology. This provides the background for a discussion of the concept of ' identity ' and of a ' theory of identity formation ' based largely but not exclusively on the work of Erik Erikson.

Then follows a chapter on methodology, and particularly the problems of studying single cases. The bulk of the book consists of a series of such cases, eight of them Indian and one Canadian. All were young students. Two were studied mainly through diaries they kept, the others rather more systematically by means of projective tests, prospective autobiographies, and interviews, repeated over a period of three years. The author presents each case briefly and then discusses it in the light of his theoretical orientation and of various conceptual issues. The cases were selected so as to be suitable for this purpose, but also so as to give, taken together, " a composite sketch of the nature of personality development during late adolescence." The Canadian case is used to provide a cross-cultural perspective. The book ends with a brief epilogue.

This approach is clearly very different from what is generally accepted as the psychological study of personality as exemplified, say, in the work of Eysenck or Cattell. It is also, I think, a good deal less sophisticated. But, in view of its relative neglect, this is only to be expected. I think the author rightly points to the potential usefulness of the study of individual cases. They certainly serve to illustrate theories; but they also serve as a check, for the question is raised, in what sense do the theories explain the individual case? I think it plausible to argue that the problem of the behaviour of individual human beings cannot be reduced to either physiology or sociology. Can one then see each individual as no more than the intersection of a set of traits? Paranjpe would certainly argue not, and support this by his cases.

The book does not set out to be a profound discussion of this and other theoretical issues. And clearly the number of general conclusions that can be derived from nine case studies is limited. Nevertheless, this is a welcome contribution: partly because it helps to redress the balance; partly because it at least raises important issues in a readily comprehensible way; and partly because the data are interesting in themselves.

I think the book will be of greatest use to students of psychology as a foundation discipline or ancillary subject: for example, in education, social work, or counselling. It could well be the basis of seminars or project work. The cross-cultural comparisons will be useful in our mixed society, though I cannot resist here a puff for that quite remarkably insightful book, *Beyond a Boundary*, by C. L. R. James (pub. Stanley Paul, 1963, 1976), perhaps little known to psychologists. The present author would be among the first to commend the use of such autobiographical accounts in understanding what it means to grow up in another culture.

JOHN RADFORD.

RUTTER, M. (1975). *Helping Troubled Children.* Harmondsworth: Penguin, pp. 368, £1·00.

Around the activities of psychiatrists and child psychiatrists a great deal of mysticism and secrecy has existed and continues to exist. Michael Rutter, in this book, sets out to dispel this. This is associated in his view with much befogging jargon, speculation and vague generalisations; all this he largely succeeds in avoiding. In approaching his task he has the general reader and the non-psychiatric professional, who cares for children, in mind. The first chapter describes how the assessment is carried out, and the disorder, if present, identified. The next four chapters outline the background knowledge necessary for such an assessment. The commoner psychiatric disorders of childhood are then discussed, and in a final chapter the emphasis is on the range and appropriateness of treatments now available.

His approach, therefore, is a well-ordered one, and this is particularly evident in his approach to assessment in chapter one. The process, as portrayed, is a highly structured one starting with a decision as to whether an abnormality exists or not, and then deciding whether it leads to any form of impairment. To help make these judgments, sets of useful criteria are given and these are followed by diagnosis. Here some useful points are made about the nature of psychiatric illness, the W.H.O. multi-dimensional approach is advocated and the principles and rationale of classification discussed. In making his assessments, he recommends that the clinician should derive hypotheses about the psychological processes involved on the basis of a functional analysis of observable behaviour; but before doing this the clinician needs to use all his powers of empathy, intuition and understanding in deciding what to look for. This is a valuable chapter in which useful distinctions are made, but it is perhaps a little misleading to give the impression that this systematic approach is typical of clinical practice.

Chapters two and three are less satisfying. They deal with child development and individual differences respectively in a total of 80 pages—an impossible task which leads to an odd assortment of sub-sections. By contrast, chapters four and five are much more convincing. Here Rutter is able to make good use of his research experience and wide reading in developing his points. Incidentally these chapters, in particular, refer constantly to a valuable list of references and lead him to admit, as he does on a number of occasions, that our understanding of various processes is imperfect. Few psychiatrists would admit this in public.

While, in discussing the commoner psychiatric disorders, he makes interesting points, his highly scientific and questioning attitude slips a little in places, for example, when illustrating his approach to treatment by means of case histories. Here he writes with a conviction that is not in keeping with his earlier attitudes, and mechanisms are listed in a manner suggesting that their existence is commonly acknowledged. This is emphasised in his final chapter on treatment, where, at the outset, he rightly points out that the various kinds of treatment have regrettably not been properly evaluated, and that it is only recently that such evaluation has been seen by psychiatrists to be necessary.

Few books exist which attempt to introduce child psychiatry, and those that do exist are usually written from a particular theoretical perspective. This book, therefore, provides a welcome addition to this limited literature, not only because Michael Rutter is admirably qualified to write such a book, but also because he is prepared to forego a particular theoretical perspective and adopt a more pragmatic outlook. The claim that one kind of therapy can cope with all childhood illnesses is rejected, and treatments are chosen, not only after a diagnosis of the illness but ideally also taking into account the personal qualities of both patient and therapist. There is some unevenness in his treatment of some topics, but this criticism probably arises from the fact that in a small book he has attempted the impossible. This calls into question its suitability for the general reader, but, for a thoughtful and perceptive person with some background in the area, it provides an excellent introduction, posing interesting, relevant and incisive questions.

R. G. A. HANNA.

PUBLICATIONS RECEIVED

Books for review have been selected from the following. (" n.p." means " no price stated.")

Anderson, J. R. (1976). *Language, Memory and Thought.* Hillsdale, New Jersey: Lawrence Erlbaum Associates, Distributor John Wiley, pp. xiii+546, £15·40, $26·50.

Baddeley, A. (1976). *The Psychology of Memory.* New York and London: Harper and Row, pp. xvii+430, p. £4·95, c. £9·75.

Bassan, V. J. (1976). *Comment Intéresser l'Enfant à l'Ecole.* Paris: Presses Universitaires de France, pp. 195, n.p.

Beard, M. (1976). *Computer Assisted Instruction: The Best of ERIC* 1973—*May,* 1966. Stanford, California: The ERIC Clearinghouse on Information Resources, pp. 38, $2·25.

Beech, H. R. (Ed.) (1976). *Obsessional States.* London: Methuen, pp. viii+352, £3·60.

Bennis, W. G., Benne, K. D., Chin, R., and Corey, K. E. (Eds.) (1976). *The Planning of Change* (third edit.). New York: Holt, Rinehart and Winston, pp. ix+517, £7·25.

Blumberg, A., and Golembiewski, R. T. (1976). *Learning and Change in Groups.* Harmondsworth: Penguin, pp. 208, £0·80.

Bourne, L. E., and Ekstrand, B. R. (1976). *Psychology: Its Principles and Meanings* (second edit.). New York: Holt, Rinehart and Winston, pp. xi+540, £9·00.

Brown, J. W. (Ed.) (1976). *Nonprint Media Information Networking: Status and Potentials.* Stanford, California: The ERIC Clearinghouse on Information Resources, pp. 60, $2·75.

Cairns, H. S., and Cairns, C. E. (1976). *Psycholinguistics: A Cognitive View of Language.* New York: Holt, Rinehart and Winston, pp. xi+252, £5·25.

Catterall, C. D. (Ed.) (1976). *Psychology in the Schools in International Perspective Volume* 1. Columbus, Ohio: International School Psychology Steering Committee, pp. iii+253, n.p.

Children's Rights Workshop (1976). *Sexism in Children's Books: Facts, Figures and Guidelines.* London: Writers and Readers Publishing Co-operative, pp. 56, £0·60.

Churchill, G. A. (1976). *Marketing Research: Methodological Foundations.* Hinsdale, Illinois: The Dryden Press, pp. xix+683, n.p.

Coombs, D. H. (1976). *Simulation and Gaming: The Best of ERIC.* Stanford, California: The ERIC Clearinghouse on Information Resources, pp. 24, $2·25.

DeCharms, R. (1976). *Enhancing Motivation: Change in the Classroom.* New York: Irvington Publishers, Distributor John Wiley, pp. ix+279, £11·55, $19·60.

Duke, C. (1976). *Australian Perspectives on Lifelong Education.* Victoria, Australia: A.C.E.R., pp. 104, $3·50.

Ellis, T., McColgan, D., McWhirter, J., and Haddow, B. (1976). *William Tyndale: The Teachers' Story.* London: Writers and Readers Publishing Co-operative, pp. 171, £1·00.

Endler, N. S., Boulter, L. R., and Osser, H. (Eds.) (1976). *Contemporary Issues in Developmental Psychology* (second edit.). New York: Holt, Rinehart and Winston, pp. xiii+689, £10·00.

Estes, W. K. (Ed.) (1976). *Handbook of Learning and Cognitive Processes, Volume* 3 : *Approaches to Human Learning and Motivation.* Hillsdale, New Jersey: Lawrence Erlbaum Associates, Distributor John Wiley, pp.x+373, £14·25, $23·50.

Eysenck, H. J., and Wilson, G. D. (Eds.) (1976). *The Experimental Study of Freudian Theories.* London: Methuen, pp. xv+405, £3·90.

Ferguson, E. D. (1976). *Motivation: An Experimental Approach.* New York: Holt, Rinehart and Winston, pp. xvi+453, £9·00.

Files, P. (1976). *A Guide to Educational Resources* 1976-77. Stanford, California: The ERIC Clearinghouse on Information Resources, pp. 38, $3·50.

Fitzgerald, R. T., Musgrave, P. W., and Pettit, D. W. (1976). *Participation in Schools? Five Case Studies.* Victoria, Australia: A.C.E.R., pp. 202, $10·00.

Frances, R., Roubertoux, P., and Denis, M. (1976). *Culture Artistique et Enseignement Superieur*. Paris: Mouton, pp. 222, 58 f.f.

Frost, J. L., and Kissinger, J. B. (1976). *The Young Child and the Educative Process*. New York: Holt, Rinehart and Winston, pp. viii+390, £8·00.

Goldfried, M. R., and Davison, G. C. (1976). *Clinical Behavior Therapy*. New York: Holt, Rinehart and Winston, pp.xii+301, £9·00.

Helms, D. B., and Turner, J. S. (1976). *Exploring Child Behavior*. Philadelphia: W. B. Saunders Co., pp.xvii+726, £14·95.

Holland, J. G., Solomon, C., Doran, J., and Frezza, D. A. (1976). *The Analysis of Behavior in Planning Instruction*. Reading, Mass., and London: Addison-Wesley Publishing, pp. xii+450, £9·10.

Hornsby, B., and Shear, F. (1976). *Alpha to Omega: The A-Z of Teaching Reading, Writing and Spelling* (second edit.). London: Heinemann, pp. x+286, £2·75.

Hornsby, B., and Shear, F. (1976). *Alpha to Omega Flash Cards*. London: Heinemann, £5·25.

Jensen, A. R. (1976). *Educability and Group Differences*. London: Methuen, pp. xiii+407, £3·20.

Kissen, M. (Ed.) (1976). *From Group Dynamics to Group Psychoanalysis*. Washington and London: Halstead Press, Distributor John Wiley, pp. xv+362, £15·50, $28·50.

Klahr, D. (Ed.) (1976). *Cognition and Instruction*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Distributor John Wiley, pp. xiii+361, £14·50, $24·55.

Klahr, D., and Wallace, J. G. (1976). *Cognitive Development: An Information-Processing View*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Distributor John Wiley, pp. xii+244, £9·40, $15·95.

Koehler, J. W., Anatol, K. W. E., and Applebaum, R. L. (1976). *Organisational Communication: Behavioral Perspectives*. New York: Holt, Rinehart and Winston, pp.xii+274, £6·00.

Kogan, N. (1976). *Cognitive Styles in Infancy and Early Childhood*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Distributor John Wiley, pp. xiii+146, £9·75, $16·50.

Krumboltz, J. D., and Thoresen, C. E. (Eds.) (1976). *Counselling Methods*. New York: Holt, Rinehart and Winston, pp. xvi+576, £8·50.

Lana, R. E. (1976). *The Foundations of Psychological Theory*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Distributor John Wiley, pp. xii+177, £9·20, $16·75.

Lickona, T. (Ed.) (1976). *Moral Development and Behaviour: Theory, Research and Social Issues*. New York: Holt, Rinehart and Winston, pp. xiv+430, £9·25.

Lipsitt, L. P. (Ed.) (1976). *Developmental Psychobiology : The Significance of Infancy*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Distributor John Wiley, pp.x+143, £7·65, $12·65.

Lomax, D. E. (Ed.) (1976). *European Perspectives in Teacher Education*. London: John Wiley, pp. x+290, £9·50, $20·00.

Mayhew, L. B. (1976). *How Colleges Change: Approaches to Academic Reform*. Stanford, California: The ERIC Clearinghouse on Information Resources, pp. 52, $1·50.

Mitchell, A. M. (1976). *The Use of Media in Career Education*. Stanford, California: The ERIC Clearinghouse on Information Resources, pp. 59, $3·25.

Norman, D. A. (1976). *Memory and Attention: An Introduction to Human Information Processing* (second edit.). New York : John Wiley, pp. xiii+262, p. £3·95, $7·20, c. £8·45, $15·45.

Oakley, A. (1976). *Housewife*. Harmondsworth: Penguin, pp. x+273, £0·80.

Ozolins, U., and Keeves, J. P. (1975). *Concerning Leisure*. Victoria, Australia: A.C.E.R., pp. vi+142, $3·50.

Pask, G. (1975). *Conversation, Cognition and Learning: A Cybernetic Theory and Methodology*. Amsterdam: Elsevier, pp. xi+570, Dfl. 96·00.

Rae, G., and McPhillimy, W. N. (1976). *Learning in the Primary School: A Systematic Approach*. London: Hodder and Stoughton, pp. 126, p. £1·50, c. £3·25.

Read, K. H. (1976). *The Nursery School* (sixth edit.). Philadelphia: W. B. Saunders, pp. viii+394, £7·00.

Reid, S. T. (1976). *Crime and Criminology.* Hinsdale, Illinois: The Dryden Press, pp. xvii+740, £9·25.

Resnick, L. B. (Ed.) (1976). *The Nature of Intelligence.* Hillsdale, New Jersey: Lawrence Erlbaum Associates, Distributor John Wiley, pp. x+364, £13·80, $22·00.

Richards, A. (Ed.) (1976). *The Pelican Freud Library Volume* 6: *Jokes and their Relation to the Unconscious—Sigmund Freud.* Harmondsworth: Penguin, pp. 330, £1·00.

Roberts, T. (Ed.) (1976). *The Circumstance of Learning.* University of Manchester: Faculty of Education, pp. 99, £1·50.

Rosenthal, R. (1976). *Experimenter Effects in Behavioural Research* (enlarged edit.). New York : Irvington Publishers, Distributor John Wiley, pp. xiii+500, £11·50 $19·50.

Saville, P., and Blinkhorn, S. (1976). *Undergraduate Personality by Factored Scales.* Slough: N.F.E.R., pp. 176, £6·50.

Sayre, K. M. (1976). *Cybernetics and the Philosophy of Mind.* London and Henley: Routledge and Kegan Paul, pp. xiii+265, £6·95.

Scottish Council for Research in Education (1976). *Burt Word Reading Test,* 1974 *Revision Manual.* London: Hodder and Stoughton, pp. vi+18, £1·00 per 20 copies (Test), £0·75 (Manual).

Seibert, W. F. (1976). *Instructional Television*: *The Best of ERIC.* Stanford, California: The ERIC Clearinghouse on Information Resources, pp. 31, $2·25.

Sellitz, C., Wrightsman, L. S., and Cook, S. W. (1976). *Research Methods in Social Relations* (third edit.). New York: Holt, Rinehart and Winston, pp. xv+624, £8·50.

Sheridan, C. L. (1976). *Fundamentals of Experimental Psychology* (second edit.). New York: Holt, Rinehart and Winston, pp. xx+475, £9·75.

Slater, P. (Ed.) (1976). *Explorations of Intrapersonal Space Volume* 1: *The Measurement of Intrapersonal Space by Grid Technique.* London: John Wiley, pp. 258, £8·75, $16·65.

Smith, D. E. P. (1976). *A Technology of Reading and Writing, Volume* 1. New York : Academic Press, pp. xi+163, £5·50.

Stewart, P. R. (1976). *Children in Distress*: *American and English Perspectives.* Beverley Hills and London: Sage Publications, pp. 285, p. £4·95, c. £8·00.

Stuart, F. (1976). *Recreation for the Retarded.* London: N.S.M.H.C., pp. 23, £0·50.

Taylor, I. (1976). *Introduction to Psycholinguistics.* New York: Holt, Rinehart and Winston, pp. xiii+434, £8·25.

Tomlinson, D. (1976). *Finance for Education in Australia*: *Developments* 1969-75. Victoria, Australia: A.C.E.R., pp. 70, $2·50.

Van Til, W. (Ed.) (1976). *Issues in Secondary Education.* Chicago, Illinois: National Society for the Study of Education, Distributor University of Chicago Press, pp. xiii+350, £8·20.

White, R. W. (1976). *The Enterprise of Living*: *A View of Personal Growth* (second edit.). New York: Holt, Rinehart and Winston, pp. xi+563, £9·75.

Whittaker, J. O. (1976). *Introduction to Psychology* (third edit.). Philadelphia and London: W. B. Saunders, pp. xvi+823, £10·00.

Williams, J. D. (Ed.) (1976). *Anthology of Ideas for Catering Education.* London: L. V. Catering Education Research Institute, Ealing Technical College, pp.v+66, £1·00.

Wilson, A. F. (1975). *The Effects of Schools in Victoria on the Science Achievement of Junior Secondary Students.* Victoria, Australia: A.C.E.R., pp. iv+48, $1·50.

Wolff, J. G. (1976). *Language, Brain and Hearing.* London: Methuen, pp. xiv+237, £3·20.

Zax, M., and Cowen, E. L. (1976). *Abnormal Psychology*: *Changing Conceptions* (second edit.). New York: Holt, Rinehart and Winston, pp. xii+643, £10·00.

THE MOST WIDELY ACCLAIMED REFERENCE WORKS IN EDUCATION AND PSYCHOLOGY

Publications of

The Institute of Mental Measurements

Edited by
OSCAR KRISEN BUROS

Tests in Print II

An invaluable descriptive bibliography of 2467 commercially available tests, and a master index to the entire *Mental Measurements Yearbook* series. Contains over 16,000 individual test references and comprehensive indexing.

1156 two-column pages
Dec., 1974; $70 per copy

Seventh Mental Measurements Yearbook

Detailed descriptions of 1157 new and revised tests with original and excerpted reviews by testing and subject area specialists, over 12,000 references, and a bibliography of books on testing.

2032 two-column pages
1972; $70 per two-vol. set

Complete Coverage of Tests, Reviews, and References in Specific Areas

PERSONALITY TESTS AND REVIEWS II	1975, 872 two-column pages, $45.
INTELLIGENCE TESTS AND REVIEWS	1975, 1158 two-column pages, $55.
VOCATIONAL TESTS AND REVIEWS	1975, 1114 two-column pages, $55.
READING TESTS AND REVIEWS II	1975, 285 two-column pages, $20.

No postage charges on prepaid orders.

Order directly from

THE GRYPHON PRESS
220M Montgomery St., Highland Park, N.J. 08904

Write for FREE Catalog of Institute of Mental Measurements Publications

PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

available from the Society's offices

St. Andrews House, 48 Princess Road East, Leicester, LE1 7DR

Psychology in Great Britain (Bulletin Supplement). Edited by Prof. B. Foss	25p
Psychology and Psychologists: A Pamphlet for Schools	15p
Teaching Educational Psychology in Training Colleges	25p
Technical Recommendations for Psychological and Educational Tests	10p
Aspects of Autism. Edited by P. J. Mittler	50p
The Psychological Study of Deafness and Hearing Impairment. Edited by M. Rodda	40p
Postgraduate Qualifications and Courses in Psychology 1976	40p

PRESIDENTIAL ADDRESSES

1965 Perceptual Defence and the Engineering Psychologist. Dr. Donald Broadbent	10p
1968 Dissonance and Stress in the Teaching of Psychology: Some Reflections of an Accidental Psychologist: Professor George Seth	10p
1969 Changing Horizons: An Essay in Autobiography: Dr. Boris Semeonoff	10p
1970 Choosing: Professor R. J. Audley	10p
1971 In Search of an Idiographic Psychology. By Professor H. Gwynne Jones	10p
1972 Psychology today and tomorrow. By Professor Harry Kay	10p
1973 Psychology in Society: Ends or End. By Professor Max Hamilton	10p
1974 On Taking Sides: Professor B. M. Foss	10p
1975 Thought and the Brain: Professor O. L. Zangwill	10p
1976 Psychology and Social Policy: Professor Jack Tizard	10p

C. S. MYERS LECTURES

Remembering Dr. Myers: Sir Frederic Bartlett	10p
Beyond Industrial Psychology: Dr. R. Conrad	10p
Practice at Liking: A Study in Experimental Aesthetics: Professor D. W. Harding	10p
C. S. Myers in Retrospect: Professor Alec Rodger	10p
On Applied Psychology in Britain: Professor G. C. Drew	10p
Measuring Psychiatric Morbidity in a South Indian Population: Professor G. M. Carstairs	10p

MEMORANDA OF EVIDENCE

To the Royal Commission on the Law Relating to Mental Illness and Mental Deficiency 1954-1957	10p
To the Royal Commission on the Penal System in England and Wales	10p
To the Royal Commission on Medical Education	10p
To the Ministry of Health and Department of Education and Science on the Responsibility for the Education of Mentally Handicapped Children	10p
To the Advisory Council on the Penal System (Sub-committee on Detention Centres)	10p
To the Departmental Committee on the Adoption of Children	10p
To the Department of Health and Social Security Sub-committee on the Role of Psychologists in the Health Services	10p
To the Butler Committee on the Law relating to the Mentally Abnormal Offender	10p
To the Warnock Committee on the Enquiry into Special Education (Summary)	10p

MISCELLANEOUS

Principles governing the employment of psychological tests and clinical instruments	7p
Psychological Tests: A Statement by the British Psychological Society	7p
Memorandum on the Establishment and Accommodation of Psychologists in General Hospitals	7p
The School Psychological Service (England and Wales)	7p
The changing role of the Educational Psychologist in Child Guidance Services in Scotland	10p
Teaching Psychology to Nurses (DCP Report)	10p
Training in Psychotherapy (DCP Report)	10p
Guidelines for the In-service Training of Probationer Clinical Psychologists	10p

(All prices on the above publications include postage)

INTERNATIONAL JOURNAL OF PSYCHOLOGY
JOURNAL INTERNATIONAL DE PSYCHOLOGIE

Published quarterly by the International Union of Psychological Science

The International Journal of Psychology is devoted to cross-cultural comparisons of psychological phenomena in general psychology, child psychology and social psychology. It also deals with international confrontation of theories and methods on which psychologists from various countries differ because of their native culture and training. Oriented towards basic research and theory, it publishes scientific papers in English or in French. In each issue, an International Platform gives news and informations on internationally significant problems.

Editor (Scientific papers)
J. Leroux
100, Tiensestraat
3000 Louvain, BELGIUM

Associate Editor (International Platform)
W. Holtzman
University of Texas, Austin, U.S.A.

1974 VOLUME 9 No. 3

CONTENTS **SOMMAIRE**

Topological and Euclidean spatial features noted by children : A cross-cultural study
GUSTAV JAHODA, JAN B. DEREGOWSKI and DURGANAND SINHA.... 159

Ecology, culture and psychological differentiation
JOHN W. BERRY and ROBERT C. ANNIS 173

Le développement de la conservation des longueurs de deux horizontales décalées : étude comparée d'enfants canadiens-français et rwandais
ADRIEN PINARD 195

A cross-cultural study of recognition memory for faces
J. W. SHEPHERD, J. B. DEREGOWSKI and H. D. ELLIS 205

Indian and Canadian choice behaviour in a Maximizing Difference Game and in a Game of Chicken
DAVID W. CARMENT 213

INTERNATIONAL PLATFORM FOR PSYCHOLOGISTS — TRIBUNE INTERNATIONALE DES PSYCHOLOGUES

Scientific and professional ethics and conduct for psychologists. A statement by the Executive Committee of I.U.P.S. : 225

Code déontologique pour psychologues. Une déclaration du Comité Exécutif de l'I.U.P.S. 227

Continuing concerns in cross-cultural ethics : A report
J. L. TAPP, H. C. KELMAN, H. C. TRIANDIS, L. S. WRIGHTSMAN and G.V. COELHO 231

Annual subscription (4 issues)	**$17.00**
Single issue	**$5.00**

Should be sent to : **Centrale des Revues DUNOD-GAUTHIER-VILLARS**
24-26, Boulevard de l'Hôpital 75005 PARIS

EDUCATIONAL RESEARCH

Vol. 19, No. 1 *Editor* : ALFRED YATES *November*, 1976

INNER LONDON EDUCATION AUTHORITY. Children with special difficulties: an ILEA research project.

JUNE PURVIS. The low status of adult education. Some sociological reflections.

P. D. PUMFREY and J. WARD. Adjustment from primary to secondary school.

T. DERRICK. The criticism of inferential statistics.

TESSA ROBERTS. 'Frustration level' reading in the infant school.

JOHN GRAY and DAVID SATTERLY. A chapter of errors: 'Teaching Styles and Pupil Progress' in retrospect.

ROY EVANS. The prediction of educational handicap—a longitudinal study.

JOAN FREEMAN. Developmental influences on children's perception.

BOOK REVIEWS.

The subscription is £4·75 a year (incl. postage); £6·50 Overseas (postage paid).

All orders should be sent to :

NFER Publishing Company Ltd. (Journals Department), 2 Jennings Buildings,
Thames Avenue, Windsor, Berks., SL4 1QS

The British Journal of Medical Psychology

Volume 48, Part 3, September 1975

John M. Fielding. A technique for measuring outcome in group psychotherapy
D. A. Shapiro. Some implications of psychotherapy research for clinical psychology
D. A. Shapiro, H. L. Caplan, P. D. Rohde and **J. P. Watson.** Personal Questionnaire changes and their correlates in a psychotherapeutic group
H. L. Caplan, P. D. Rohde, D. A. Shapiro and **J. P. Watson.** Some correlates of repertory grid measures used to study a psychotherapeutic group
J. H. Kahn. 'Do not interpretations belong to God?' The validity of interpretations in psychotherapy and in literature
Eleanor S. Wertheim. Person-environment interaction: the epigenesis of autonomy and competence. III. Autonomy and para-/pre-linguistic and linguistic action systems: review of developmental literature (normal development)
S. Stephanos. The object relations of the psychosomatic patient
Roy Muir. The family and the problem of internalization
Michael Humphrey. The effect of children upon the marriage relationship
C. Heijn, P. G. Myerson and **P. Schmitt.** An approach to the supervision of para-professionals working with the mentally ill
Review
Short Book List

Volume 50 (1977) £16·00 ($39·50 *U.S.A. and Canada*). *Single parts* £5·00 ($11·75 *U.S.A. and Canada*), *each.*

Enquiries about advertising in the Journal should be sent to the London or New York office of the Press.

CAMBRIDGE UNIVERSITY PRESS
BENTLEY HOUSE, 200 EUSTON ROAD, LONDON NW1 2DB
32 EAST 57TH STREET, NEW YORK, N.Y. 10022

The British Journal of Psychology

Volume 66, Part 3, August, 1975

ALAN BRANTHWAITE. Subjective value of information
DAVID FREIDES. Information complexity and cross-modal functions
ANTHONY GALE, GRAHAM SPRATT, BRUCE CHRISTIE and ADRIAN SMALLBONE. Stimulus complexity, EEG abundance gradients, and detection efficiency in a visual recognition task
MANSUR LALLJEE and MARK COOK. Anxiety and ritualized speech
ALAN GAULD. A note on inappropriate constancy-scaling and the Muller-Lyer illusion
ANGUS R. H. GELLATLY and VERNON H. GREGG. The effects of negative relatedness upon word-picture and word-word comparisons and subsequent recall
D. R. GODDEN and A. D. BADDELEY. Context-dependent memory in two natural environments: on land and underwater
JOHN T. E. RICHARDSON. Imagery and deep structure in the recall of English normalizations
RUBEN C. GUR and ERNEST R. HILGARD. Visual imagery and the discrimination of differences between altered pictures simultaneously and successively presented
PETER WALKER. The subliminal perception of movement and the ' suppression ' in binocular rivalry
PAUL J. BARBER and J. PHILIPPE RUSHTON. Experimenter bias and subliminal perception
ARIE W. KRUGLANSKI. Context, meaning and the validity of results in psychological research
D. G. BOYLE. Research note: a new technique for studying phenomenal causality
BOOK REVIEWS
OTHER PUBLICATIONS RECEIVED

Volume 68 (1977) £20·00 ($49·50 U.S.A. and Canada). Single parts £6·00 ($15·00 U.S.A. and Canada) each.

Enquiries about advertising in the Journal should be sent to the London or New York office of the Press.

CAMBRIDGE UNIVERSITY PRESS

Bentley House, 200 Euston Road, London, NW1 2DB
32 East 57th Street, New York, N.Y. 10022

THE BRITISH JOURNAL OF SOCIAL AND CLINICAL PSYCHOLOGY

Volume 14, Part 3, September, 1975

ALLAN I. TEGER and NATHAN KOGAN. Decision-making for others under reciprocal and non-reciprocal conditions
R. LYNN and S. L. HAMPSON. National differences in extraversion and neuroticism
G. M. STEPHENSON and C. J. BROTHERTON. Social progression and polarization: a study of discussion and negotiation in groups of mining supervisors
TERRY W. MALLENBY and RUTH G. MALLENBY. The personal space of hard-of-hearing children after extended contact with ' normals '
HAZEL E. NELSON and PAT McKENNA. The use of current reading ability in the assessment of dementia
MICHAEL W. EYSENCK. Arousal and speed of recall
DAVID A. WINTER. Some characteristics of schizophrenics and their parents
GEORGE SERBAN. Relationship of mental status, functioning and stress to readmission of schizophrenics
GORDON E. BARNES. Extraversion and pain
PHILIPPE J. RUSHTON and DIANE OWEN. Immediate and delayed effects of TV modelling and preaching on children's generosity
STEVEN W. DUCK. Attitude similarity and interpersonal attraction: right answers and wrong reasons
ALLAN SIGSTON and DAVID G. WHITE. Conformity in children as a function of age level
BOOK REVIEWS
OTHER PUBLICATIONS RECEIVED

Volume 16 (1977) £19·00 ($47·50 *U.S.A. and Canada*). *Single parts* £6·00 ($14·00 *U.S.A. and Canada*) *each.*

Enquiries about advertising in the Journal should be sent to the London or New York office of the Press.

CAMBRIDGE UNIVERSITY PRESS

BENTLEY HOUSE, 200 EUSTON ROAD, LONDON, NW1 2DB
32 EAST 57th STREET, NEW YORK, N.Y. 10022

PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

The British Journal of Psychology
The British Journal of Medical Psychology
Journal of Occupational Psychology
The British Journal of Social and Clinical Psychology
The British Journal of Mathematical and Statistical Psychology
The British Journal of Educational Psychology

The British Journal of Psychology, *The British Journal of Medical Psychology*, *Journal of Occupational Psychology* and *The British Journal of Social and Clinical Psychology* appear in quarterly parts. The 1977 subscription prices, payable in advance, are £20·00 (US $60·00 in the USA and Canada) for *The British Journal of Psychology*; £19·00 (US $57·00) for *The British Journal of Social and Clinical Psychology*; £16·00 (US $48·00) for *The British Journal of Medical Psychology* and £13·00 (US $39·00) for *Journal of Occupational Psychology*. Subscriptions to any of these journals may be sent to any bookseller or subscription agent, or to the Cambridge University Press, Bentley House, 200 Euston Road, London, NW1 2DB. American and Canadian enquiries should be addressed to the Cambridge University Press, American Branch, 32 East 57th Street, New York, N.Y. 10022.

The British Journal of Mathematical and Statistical Psychology is issued twice a year, in May and November. The subscription rate, payable in advance, is £15·00 per volume post free (US $34·00 in the USA). Subscriptions should be sent to the Society's agents, John Wright and Sons, Ltd., 42-44 Triangle West, Bristol, BS8 1EX, or to any bookseller.

The British Journal of Educational Psychology is issued three times a year, in February, June and November. The prepaid subscription is £8 a year, post free, or £3·00 each number. The prepaid subscription price in the USA is $20·00 post free. Subscriptions should be sent to the Scottish Academic Press, 33 Montgomery Street, Edinburgh, EH7 5JX.

The series of *Monograph Supplements* issued in connection with *The British Journal of Psychology* is now discontinued. Orders for back volumes may be sent to any bookseller, or the Cambridge University Press, Bentley House, 200 Euston Road, London, NW1, from whom details of these publications may be obtained.

MANUSCRIPTS

For *The British Journal of Psychology* should be sent to Professor A. D. B. Clarke, Department of Psychology, The University, Hull, HU6 7RX;

For *The British Journal of Medical Psychology* should be sent to Professor Arthur H. Crisp, Psychiatric Research Unit, Atkinson Morley's Hospital, Copse Hill, London, SW20 0NE;

For *Journal of Occupational Psychology* should be sent to Professor B. Shackel, Department of Human Sciences, University of Technology, Loughborough, Leicestershire.

For *The British Journal of Social and Clinical Psychology* should be sent to Dr. Halla Beloff (Social Psychology Editor), Department of Psychology, University of Edinburgh, 60 Pleasance, Edinburgh EH8 9TJ; or to Dr. H. R. Beech (Clinical Psychology Editor), Department of Psychiatry, Withington Hospital, West Didsbury, Manchester, M20 8LR;

For *The British Journal of Mathematical and Statistical Psychology* should be sent to Professor P. Levy, Department of Psychology, University of Lancaster, Lancaster, LA1 4YF;

For *The British Journal of Educational Psychology* should be sent to Professor N. J. Entwistle, Department of Educational Research, University of Lancaster, Lancaster.

Contributors are advised to consult *Suggestions to Authors* issued by the Standing Committee on Publications of the British Psychological Society and obtainable, price £0·50 (USA $2·00) post free, from Cambridge University Press.

All manuscripts should be submitted in duplicate.

OFF-PRINTS

Contributors on application receive fifty copies of their papers free. Additional copies may be obtained in units of fifty copies.

SUBSCRIPTIONS TO THE BRITISH PSYCHOLOGICAL SOCIETY

The subscription for Fellows, Associates and Graduate Members is £18·00 (£16·00 if paid by Banker's Order).

They are entitled to purchase any or all of the Society's Journals at the special rates of £4·00 per Journal per annum.

Information on membership and activities may be obtained from the Secretary, The British Psychological Society, St. Andrew's House, 48 Princess Road, Leicester, LE1 7DQ.

CONTENTS

SYMPOSIUM: PRACTICAL AND THEORETICAL ISSUES IN PIAGETIAN PSYCHOLOGY—III

SMEDSLUND, J.: Piaget's Psychology in Practice 1
BROWN, G., and DESFORGES, C. W.: Piagetian Psychology and Education: Time for Revision . 7
ANTHONY, W. S.: Activity in the Learning of Piagetian Operational Thinking 18

ARTICLES:

HARGREAVES, D. J.: Sex Roles in Divergent Thinking 25
WATSON, J. M.: The Influence of Context on Referential Description in Children 33
SINGLETON, C. H. Redundancy in the Concept Attainment of Adolescents . 40
ELLIOTT, C. D., and MURRAY, D. J. The Measurement of Speed of Problem Solving and its Relation to Children's Age and Ability 50
CONRAD, R.: Lip-Reading by Deaf and Hearing Children 60

RESEARCH NOTES:

FRANKLIN, B. S., and RICHARDS, P. N.: Effects on Children's Divergent Thinking Abilities of a Period of Direct Teaching for Divergent Production . 66
RICHARDSON, K.: Reading Attainment and Family Size: an Anomaly. . 71
WONG, B. Y. L.: Further Investigation of the Emergence of Ability for Partial and Complete Multiplication of Classes in Young School Children 76
LAWTON, J. T.: The Development of Causal and Logical Connectives in Children 81
FAIRWEATHER, H., and BUTTERWORTH, G.: The WPPSI at Four Years: A Sex Difference in Verbal-Performance Discrepancies 85
LAWSON, M. J., and JARMAN, R. F.: A Note on Jensen's Theory of Level I Ability and Recent Research on Human Memory. 91
CONFERENCE ABSTRACTS 95

CRITICAL REVIEW:

WILSON, J. A.: How Many Teaching Styles? How much Progress? . . 105
BOOK REVIEWS 108
PUBLICATIONS RECEIVED 113

UK ISSN 0007–0998

THE BRITISH
JOURNAL OF
EDUCATIONAL
PSYCHOLOGY

JUNE 1977

VOL. 47, PART 2

THE
BRITISH JOURNAL
OF
EDUCATIONAL PSYCHOLOGY

Edited by
N. J. ENTWISTLE

Assistant Editors :
S N. BENNETT
G BROWN

Sub-Editor :
S. M. CUNNINGHAM

Advised by
L. B. BIRCH
E. A. PEEL
P. E. VERNON

Vol. 47, Part 2

The British Journal of Educational Psychology is issued by the British Psychological Society once a term, in February, June and November. The subscription is £8·00 a year post free, payable in advance, or £3·00 each number.

Subscriptions and Orders should be sent to Scottish Academic Press, 33 Montgomery Street, Edinburgh, EH7 5JX. Members of the British Psychological Society receive the *Journal* on special terms.

Papers for publication should be sent to Professor N. J. Entwistle, Department of Educational Research, University of Lancaster, Lancaster. Guidance to contributors on the preparation of papers for the *Journal* is given at the end of the November number each year.

All *Books for Review* for the *British Journal of Educational Psychology* should be sent to Professor N. J. Entwistle at the above address. *Advertisements* for this *Journal* should be sent to Scottish Academic Press.

Published by
SCOTTISH ACADEMIC PRESS, 33 MONTGOMERY STREET, EDINBURGH, EH7 5JX,
for the British Journal of Educational Psychology, Ltd.

All rights reserved.

© 1977. The British Journal of Educational Psychology.

PRINTED BY BIRMINGHAM PRINTERS, LTD., 8-24 SAMPSON ROAD NORTH, BIRMINGHAM.

Br. J. educ. Psychol., 47, 117-125, 1977

SYMPOSIUM: READING ABILITIES AND DISABILITIES

CHILDREN'S STRATEGIES IN LEARNING TO READ

BY HAZEL FRANCIS
(*School of Education, University of Leeds*)

SUMMARY. Sentences taken from a particular first 'reader' in a well-known school reading scheme, together with variants of these sentences, were read by a sample of 5-year-old children whose reading skills enabled them to attempt the task. Twenty-four children, 12 of each sex, were selected from each of two very different schools using the reading scheme. The results of planned comparisons between error scores in different forms of the sentences, together with analysis of types of error showed a strong tendency to revert to familiar sentences in the face of novel versions. There were also signs of early steps in perceiving the letter/word coding of written language. Knowledge of spoken language and of the elements of a phonic approach seemed of little relevance at this very early stage of reading. Reading strategies suggested by this study have already been reported with limited samples of children and learning contexts, and this study tends to show their relative independence of both materials and methods of teaching. Implications for an understanding of early reading and for approaches to instruction are discussed.

INTRODUCTION

THE work reported in this paper was designed to explore the strategies used by 5-year-olds reading sentences from the first book in their school reading scheme, and to test predictions of error patterns derived from the work of MacKinnon (1959) and Soderbergh (1971).

MacKinnon worked with children in their first year at school. He monitored individual and group learning with an emphasis on sentence reading. He used the Gibson-Richards reading materials (Gibson and Richards, 1957), which present simple sentences in conjunction with pictures to illustrate the nouns and to cue a few other words such as *his* and *in*. The range of letters used is small, but no attempt is made to guide letter-sound correspondences; the reader is left to discover these through the repeated reading of graded, relatively small changes in the text. According to MacKinnon children tend to make characteristic initial reading errors. These can be summarised as errors due to reverting to a previous (familiar) sentence; errors due to giving a semantically acceptable version according with the illustration, as in *This hat is in a hand* for *His hat is in his hand*; and errors due to giving a semantically acceptable but 'untrue' version, as in *This is a hand* for *This is his head.* The similarity between sentence structures leads to error in interchangeable 'frame' words, while word similarity leads to misreading based on graphic appearance. In the first example given above, *a* for *his* involves the substitution of an appropriate 'frame' word which bears no graphic similarity to the original. In the case of *This* for *His* the error is more complex, for there is strong graphic similarity as well as syntactic suitability.

Soderbergh described the reading of a Swedish 2- to 3-year-old who learned from specially prepared materials. Emphasis was placed on word learning in a context of meaningful sentence reading. At first the child learned a number of words associated with everyday objects and actions, these being presented on individual cards. They could be combined to make the sort of 'content' word sentences typical of the 'telegraphic' speech of a young child.

' Frame ' or ' function ' words were little used at first since the child indicated that there were difficulties attached to them. A typical comment was that *tongue* (tunga) was easy to show but *what* (rad) was not possible. Learning was thoroughly supervised so that the child built a firm word store that enabled her to begin to comment on similarities between new and old words. This led to typical errors in attempting new words and eventually to an analytic approach which became finer-grained as the written code was broken. After a period of a year, using specially compiled diary booklets, the child was able to tackle almost any new word presented to her.

Soderbergh described two major categories of error in tackling new words—errors due to misreading and those attributable to analytic reading. Misreadings occurred when a new word was mistaken for a known word, and they seemed to be based on different degrees of graphic similarity. The earlier misreadings were frequent and depended simply on letters in common, as in *berg* for *stag* and *sta* for *sag*. Later misreadings showed letter order in common, as in *mugg* for *mum*, and *bil* for *mil*, and they varied in the extent to which the similarity held. Misreadings such as *kaka* for *haka* suggested a very close recognition of the graphic pattern. Improved pattern recognition within words was the basis of analytic reading which was described by Soderbergh as recognition of known words, morphemes or syllables in new words. Errors involved misreading of the other parts of the words or inappropriate combination of recognised elements. Analytic reading was only possible when a fairly large store of known words was available, and when useful patterns of letters had been worked out.

The two studies combine to give some idea of the kind of strategies and errors that may be encountered in early reading, but it may be asked whether such findings are limited to learning with the tightly controlled guidance and the specially graded materials adopted in them. In order to attempt something of an answer to this question an exploration of reading errors made by children from two very different schools was undertaken. Each school used the *Gay Way* reading scheme, and sentences from the first book in the series were presented in different ways to test the following predictions, derived from the above findings.

A. Differences between readings of different presentations will show:
   (i) fewer errors in familiar sentences than in any other form;
   (ii) fewer errors in other sentences containing the same words (being somewhat familiar and meaningful) than in lists of those words;
   (iii) fewer errors in familiar sentences rearranged internally to produce novel forms than in mixed sentences combining elements from different sentences to produce new orders and new meanings.

B. Types of word error will comprise:
   (i) substitution of an appropriate ' content ' or ' frame ' word;
   (ii) misreading based on graphic similarity to known word;
   (iii) error in analytic reading.

METHOD

Sample.

A particular selection of children was made to fulfil two criteria. They were in their first year at school, between 5 and 6 years of age. They had begun to read the *Gay Way* Series (Boyce, 1949) of reading books, and could read at least some of the words on the first eight pages of the first (Red) book. (A few could read most of them, and had progressed to the next reader.) Twelve boys and 12 girls were chosen from each of two schools. Differences between the schools as described below meant that in one, termed the LC school, children were excluded because they were not able to tackle the materials ; while in the other, the MC school, it was barely possible to find enough children at a sufficiently elementary level. A comparison of the two groups showed that selection on the basis of reading skill resulted in two groups with very similar mean scores on the Goodenough test of drawing ability, a test which is often used as a measure of all-round ability, or mental maturity.

Schools.

Although the two schools were similar in that they used the same reading scheme and in that the teaching was not as tightly controlled as in the MacKinnon and Soderbergh studies, they differed in several important respects. The LC school drew children from an area in which most parents followed occupations characteristic of social class categories IIIb to V, while the MC school drew on one in which most parents could be described in terms of classes IIIa to I. The approaches to the *Gay Way* reader also differed. The LC school included the early stages of a ' Break-Through to Literacy ' scheme in its introductory reading activities, thus emphasising sentence-building and recognition. A phonic approach to gaining cues to words was also encouraged. The *Gay Way* words were introduced as an extension of this approach. The MC school used a word recognition approach to reading and a flash card technique to introduce the words of the *Gay Way* book, tending to introduce the sentences only when some word recognition was established. Thus, the overall differences in learning contexts included variables of several kinds.

Materials and procedures.

Four ways of presenting material taken from the first eight pages of the *Gay Way* first reader *The Red Book*, were used. First, actual sentences were taken from the book, and these were regarded as being familiar to the children. Then these sentences were rearranged so that, although they retained very much the same meaning and contained the same words, they were novel to the children. A further set of sentences was constructed by taking parts of different book sentences and recombining them to make new ' mixed ' sentences. Lastly, two lists of words were drawn up, the first containing all the words in the familiar sentences drawn from the first page or two of the book, and the second containing those from the later pages up to the eighth. There was thus some overlap between lists. Each sentence or list was written on a separate page of a prepared booklet, and the pages were arranged so that each novel sentence immediately preceded or followed its corresponding familiar form. Two booklets were needed to balance the arrangement, and these were used alternately, each child reading from one or the other. The mixed sentences and lists followed the other sentences in each of two sections of the booklets, both booklets being identical in this respect. The actual sentences and lists in their order in one of the booklets can be seen below in the section on results.

The following procedure was adopted. Each child was seen individually and asked to read aloud from the first eight pages of *The Red Book*. Some sentences, being superfluous to the experimental materials, were by-passed. If the child was not successful at any point the correct word was supplied; and, if it proved necessary, complete sentences were read to the child with both reader and child pointing together to each word as it was read. Thus, all children were reminded of the book sentences before they attempted the test materials. Children who found reading the book impossible, or nearly so, did not continue with the experimental reading. In the case of those who did continue each child was instructed to read the prepared booklet, reading as much as possible and pointing to each word as it was pronounced. Unknown words were attempted, but if no trial could be made the child was asked to move on to the next word. Each child was told before starting that if he or she wished to be told any of the unknown words they would be supplied at the end. Only a few availed themselves of the opportunity to ask. Each child's reading was recorded on a prepared sheet, note being taken of each word as it was read, and particular care was taken over omissions. With instances of self-correction the child's preferred version was recorded.

RESULTS

A. Sentence errors.

In order to test the set of predictions about sentence errors each mode of presentation was the subject of a planned comparison. Since it was apparent that there was considerable correspondence between the patterns for errors and for omissions the two were combined. The total figures for each word in each sentence or list are shown below, the letter F denoting familiar, N novel, M mixed and L list, and the maximum possible number of errors per word being 48.

1. (F) The big red lorry went up the hill
2 4 4 2 5 1 0 1

2. (N) Up the big hill went the red lorry
4 11 8 5 16 13 5 5

3. (N) Big pots and little pans, little pots, big pans
23 18 10 6 30 12 16 8 30

4. (F) Big pans, little pans, big pots and little pots
21 29 5 26 6 16 7 3 14

5. (F) Up and up and up went the pots
3 9 4 9 4 12 8 16

6. (N) The pots went up and up and up
4 20 17 5 13 7 11 6

7. (N) Down fell a tin pot
28 22 10 24 19

8. (F) A tin pot fell down
10 25 18 23 17

9. (M) A little tin lorry went on and on
10 5 28 7 13 22 15 22

10. (M) A big pot and a red pan fell down the hill
10 6 25 10 7 19 31 19 11 9 6

L (1) lorry pots red hill a down pans big fell tin up little
11 23 8 5 6 15 34 6 18 33 4 5
the went and
11 17 15

11. (F) Meg the hen went down the hill
0 4 2 3 6 2 1

12. (N) Down the hill went the hen, Meg
23 9 3 13 13 23 10

13. (N) The little house is a tin pot to live in
8 5 11 10 3 19 10 23 18 17

14. (F) The tin pot is a little house to live in
8 20 13 8 3 3 7 15 13 13

15. (M) Meg the hen had a tin pot to live in
1 1 2 30 8 21 16 16 10 10

L (2) house live is tin Meg pot in had to hill little a went
8 22 11 26 2 21 18 36 19 9 5 8 13
hen the down
29 14 16

Each child was given a score for each mode based on the proportion of omissions and errors to the total number of words. The mean scores for each school in each mode are shown in Table 1. It will be noticed that each school group is subdivided into the better, the middle-range and the poorer readers on the basis of their total reading score over all modes.

TABLE 1

MEAN ERROR SCORES FOR DIFFERENT VERSIONS OF READING MATERIAL.

Level of reading ability	School LC (N=24)				School MC (N=24)			
	Familiar	Novel	Mixed	List	Familiar	Novel	Mixed	List
High........	·016	·058	·082	·117	·029	·053	·056	·077
Medium	·130	·239	·271	·230	·138	·202	·185	·258
Low	·543	·646	·638	·730	·346	·566	·548	·492
All levels	·230	·315	·330	·359	·171	·274	·263	·276
Schools combined..	Familiar ·200	Novel ·294	Mixed ·297	List ·317				

The results show that although the level of error score differs for the two schools, the pattern is very similar. At all levels of reading ability the familiar sentences are markedly more error-free. This accords with the first prediction. There is a trend towards greater error in the lists than in other sentences than the familiar; so the second prediction tends to be borne out, but much less certainly. There is no tendency for the novel sentences to yield less error than the mixed, so the last prediction is not sustained. The summary scores for the two schools together make a clear statement of the results.

If the selected schools could be regarded as a random sampling of schools using the *Gay Way* scheme (and given the complexity of the differences between them there were no grounds for treating them as any particular fixed effect), then the probability of obtaining the recorded differences between error scores in the different modes of presentation by chance would be less than 0·01 (see Table 2). Such a finding suggests that the pattern of differences might obtain generally amongst children using this scheme, and that the strategies reported by MacKinnon can be found more widely than in a particular school and with other materials than the Gibson-Richards.

TABLE 2

ANALYSIS OF VARIANCE OF ERROR SCORES, SCHOOLS BEING TAKEN AS A RANDOM EFFECT WITH OTHER EFFECTS FIXED.

Source	df	MS	F	P<
Materials (M)	3	·131	32·75	·01
Schools (S)	1	·188	20·89	·01
Levels (L)	2	4·276	45·01	·01
M×S	3	·004	·44	—
M×L	6	·013	·81	—
S×L	2	·095	10·55	·01
M×S×L	6	·016	1·77	—
Within cells	168	·009	—	—
Planned comparisons:				
Familiar/others	1	·377	41·88	·01
List/novel+mixed	1	·015	1·66	—
Novel/mixed	1	·000	—	—

It can be concluded that at this early stage of reading children are helped by *familiar sentence contexts*, but that other sentences function little better than lists. Children in this study were helped to give a correct rendering when they recognised aspects of sentences they had already encountered. Their strategy was attempted recall. Novel sentences could not at first be tackled, but as the individual words of familiar sentences became well learned then new sentences incorporating these, or aspects of these, were attempted. Since not all the familiar sentences were equally well managed it was possible to explore how this came about. Sentences 1 and 11 were comparatively easy, but 4, 5, 8 and 14 yielded more error. A comparison of the well-learned with the less well-learned was made, omitting sentences 4 and 5 because they contained so much repetition. The well-learned and less well-learned yielded no differences in the error pattern of familiar *versus* novel forms of the sentences; but the well-learned yielded more list error and the less well-learned yielded less list error than could be expected on the basis of equal probabilities (χ^2=42·2, df=1, P<0·01). This finding supports the subjective impression that some of the familiar sentence reading was cued by some feature of the sentence, but that each individual word was not necessarily read.

Certain errors indicated even more clearly that some of the reading was based on only a fraction of the actual text, and that, as MacKinnon suggested, it showed a tendency to revert to a familiar form. Such errors occurred in the novel and mixed sentences. An extreme example was the reading of an entirely different sentence from the one given, when *Down the hill went the hen, Meg* was read as *Meg the hen had no house to live in.* The cue seemed to be the preceding familiar sentence which also preceded the mistaken version in the *Red Book* itself. The child's mistaken rendering was thus another familiar sentence. Less extreme examples involved the substitution of a phrase which made the novel or mixed form more like a sentence from the book, as shown in the following cases:

6. (N) The pots went up and up and up
Reading: The pots went up the hill
Familiar in book: The pots and pans went up the hill in the lorry

9. (M) A little tin lorry went on and on
Reading: — big — lorry went up the hill
Familiar in book: The big red lorry went up the hill.

15. (M) Meg the hen had a tin pot to live in
Reading: Meg the hen went up the hill
Familiar in book: Meg the hen went down the hill
Also,
Reading: Meg the hen had no house to live in
Familiar in book: Meg the hen had no house
Jip had no house to live in

It was interesting to note that mistakes rested within the domain of the school reading book. There was no tendency to fill in with words or phrases from everyday speech.

B. Word errors.

Errors which consisted of single word misreadings in otherwise correct phrases, and those which occurred when a sentence was reduced to a string of words rather than a meaningful frame were classified according to the predicted types of error. It was obvious that some items could be doubly classified, as, for example, when a substitute word was both semantically appropriate and also graphically similar to the text word. This was no problem, however, since the aim was to check whether all the predicted categories were required and whether they were adequate.

It was found that substitution of both 'content' and 'frame' words occurred as predicted, yielding both grammatical and sensible sentences. Again the substitutions were derived from the book vocabulary. An additional, unpredicted category seemed to be required for those substitutions which were appropriate following that part of a sentence already read, but which led to error and confusion thereafter. Perhaps MacKinnon and Soderbergh used such short, tightly-controlled sentences that this rarely if ever occurred; but, given the *Gay Way* materials, the children were obviously in difficulty when unable to anticipate or read on to the sentence end.

Misreadings which showed some graphic similarity with the correct version were quite frequent, especially if one letter in common was accepted as a criterion. Few misreadings showed the common order patterns described by Soderbergh, but in view of the limited materials used and the even more limited word knowledge of most of the children, this was hardly surprising. A special case of similarity was the capital letter, since the form of print seemed to remind children of other words in the book beginning with capitals whether there was any letter similarity or not.

The category of analytic reading, although predicted, proved to have few candidates for entry. Possibly *in* for *tin*, *we* for *went* and *a* for *and* may have been examples, but the opportunity for analytic readings was really very slight.

A further unpredicted category seemed to be required, however, for a rather unusual set of mistakes. Instead of the required word children sometimes gave a word which lay adjacent to it in the book text. It seemed possible that in reading their book children had not always clearly distinguished each word, and that while, for example, *fell down* was read aloud correctly in a familiar sentence, it was not clear to the child that *fell* was *fell* and *down* was *down*. Renderings of *lorry* and *down* for *went*, *red* for *lorry*, and *and* for *the* seemed to be further examples of this kind of error.

When both the predicted and the additional categories had been considered some errors seemed to fit none, but, given the flimsy associations already described as bases for error, they may well have been the consequences of something more than random guessing. At the very least they stayed within the vocabulary of the reading book.

DISCUSSION

It would seem that the reading strategies reported by MacKinnon and Soderbergh can also be seen, at least to some extent, in the early reading of 5-year-olds in more varied conditions of learning and with less tight control of materials and of the learning process. Whether any early reading material can be found in which such strategies are not used is still an open question, but, in spite of considerable differences between schools and between teaching methods, the children in this study seemed to try to read in much the same way. They produced the same patterns and types of error, with no interaction between the variations in the reading task and the effects of either school or reading ability differences. They did not draw on language outside that of the ' reader '; and in both schools some children read fluently, some read word by word, and some read only a selection of words. None appeared to attempt to ' sound ' a word they could not recognise on sight, most being too unsure of the graphic patterning of new words to do so, either intuitively or as a response to the kind of interest in phonic word-building which was being introduced in the LC school. While misreadings which bore letter similarities to the actual word were common some of the more ' advanced ' categories of graphic mis-readings and the category of analytic reading error suggested by Soderbergh were scarcely applicable to the errors in this study. It appeared that the children were not sufficiently advanced along the path of letter pattern analysis; and it was obvious that this was generally due to insufficient word learning, although individual differences between children suggested that speed of learning to discriminate and recognise patterns might vary considerably.

To summarise the children's approach to reading as shown by this study it would be fair to say that they seemed to learn by storing the paired written and spoken versions of a set of words, or a sentence or two, which formed the ground for their further attempts to learn. Then new learning was not simply the addition of more words, but their incorporation into what was developing as a structured form of knowledge of the graphic coding of written language. Using any available context clues, including expectations about the general coverage of what was to be read, they attempted to identify new words or sentences by calling on their memories of aspects of the visual appearance of what they knew. A few children were showing signs of the more developed pattern recognition and analytic skills described by Soderbergh as the precursors of the stage of ready reading. With the exception of those few who could yet only repeat a few well-learned words, the children were essentially approaching reading as a problem-solving task; and even the apparent rote learning of the exceptions may well have concealed some problem solutions.

In order to understand early reading, it is worth spelling out what this problem-solving approach involves. It is clear from the data in this study that a word may seem to be known in one context but not in another. Successful reading in generally familiar material is easier than isolated word recognition or reading in a new context. Word memory is not a simple all or nothing, but a more or less successful reconstruction from visual form and context. It is interesting that children in the MacKinnon study tended to say at first, *I don't know that word,* but later to say, *I don't remember it.* Given that word learning shows varying degrees of solidity there is obviously a need for practice of various kinds to consolidate early learning sufficiently for it to be a firm basis for attempts to make valid comparisons and discriminations that will serve to help read new words. Moreover, the actual content of this basic learning should be sufficiently broad and patterned to allow visual comparisons that can lead to a sense of system and thus to constructive attacks on new or inadequately recalled words. This may not be so much a matter of defining some ideal basic set of words as of making sure that each child develops a sound personal store. How much guidance the children need to make constructive analyses is not clear, nor is it clear how far their efforts are based on conscious rather than intuitive processes, but no child can succeed with an inadequate basic knowledge.

Children may vary in their rate of learning a basic store, but, whether they be quick or slow, this grasp is not independent of some knowledge of what written language amounts to, for the visual form of words must be invested with some meaning to be memorable. This is, at the very least, a complex meshing of the association of written words with referent pictures or objects and with their spoken forms, and of the association of written words with each other in meaningful sentences. In view of the limited range of words at first encountered, and the importance of seeing relations within the text, comparatively little demand may be made at first on the child's spoken language skills. Until children can fairly readily read a range of words they are limited in their reading strategies to using whatever visual cues they can identify to prompt recall of some sentence they know they have previously encountered *within a reading experience.* Only later do sentence 'frames' and 'content' word meanings cue more varied sentence structures and meanings. But although the usefulness of speech may at first be limited virtually to the investing of individual words with some meaning, familiarity with the use of written signs and symbols, and of strings of these as meaningful representations, may be crucial for success. A subjective impression of the difficulties of the slow readers, particularly in the LC school, was that, whether or not they found memorising and comparison difficult, they found the nature of print and of reading it virtually incomprehensible. Thus, their attempts to learn seemed to be meaningless memorising—a self-defeating enterprise.

The burden of this discussion has been that children seem to bring to initial reading those analytic, mostly intuitive, problem-solving skills that characterise human learning in general. Learning to read produces natural errors as a by-product of the process, as does early spoken language learning, and these errors indicate the aspects of coding that are being tackled and the ways they are being attempted. They can, therefore, be used to guide instruction and the selection of reading material. While linguistic descriptions of codings have influenced the design and choice of reading schemes, general psychological considerations may have been relatively neglected. Perhaps this paper has helped to describe those more universal features of strategies of learning that teachers may foster whatever the materials and methods they employ.

REFERENCES

Boyce, E. R. (1949). *The Gay Way Series: The Red Book.* London: MacMillan Education.
Gibson, C. M., and Richards, I. A. (1957). *First Steps in Reading English.* New York: Pocket Books.
Mackinnon, A. R. (1959). *How Do Children Learn to Read?* Vancouver: Copp Clark.
Soderbergh, R. (1971). *Reading in Early Childhood.* Stockholm: Almqvist and Wiksell.

(*Manuscript received 25th October,* 1976)

Br. J. educ. Psychol., 47, 126-131, 1977

READING RETARDATION AND CROSS-LATERALITY IN RELATION TO SHORT-TERM INFORMATION PROCESSING TASKS

By T. J. WHEELER, E. J. WATKINS, and S. P. McLAUGHLIN
(*Department of Communication Studies, Sheffield City Polytechnic*)

Summary. The performance of 10 cross-lateral children was compared with a control group of 10 uni-lateral children on short-term information processing tasks using three types of units—digits, letters and symbols. The tasks consisted of presenting different numbers of units of information simultaneously by tachistoscopic exposure. The cross-lateral group took significantly longer to identify accurately the information presented : their performance significantly deteriorated as the information load was increased from three to five units. The cross-lateral group was also significantly inferior to the matched control group on reading age. It is contended that the reading retardation associated with cross-laterality (which is indicative of ill-established cerebral dominance) is a specific manifestation of a general limitation in any kind of information processing.

INTRODUCTION

Numerous previous studies have shown that retardation in reading is related in some way to laterality (Gates and Bond, 1936). Smith (1950) found that retardation in reading was related to ill-established cerebral dominance, whereas Hillman (1956) questioned whether isolated factors such as handedness and eyedness were, in fact, good predictors of reading ability. However, cross-laterality specifically has been shown by Harris (1957) to become a much more reliable indicator of reading failure as a child gets older. This is consistent with Zangwill (1962) who states that children with ill-established cerebral dominance (cross-laterality) may be suffering either a maturational lag or minimal brain damage and consequently as some of the subjects mature, one is left with children in the latter category. Belmont and Birch (1965), however, conclude that there is no relationship between lateral dominance and disability in reading. This is a problematic conclusion not only because it conflicts with other findings but because there are many reasons why children may fail to read: emotional disturbance, lack of early schooling or uncorrected hearing and sight difficulties, as well as problems of different varieties of laterality (Bannatyne, 1971) and these reasons have not been clearly differentiated.

The major problem then with these studies is the lack of homogeneity of the groups of non-readers investigated. The resulting inconsistency of findings is thus not surprising. Yet, if one takes a group of children who have problems of cerebral dominance which have been precisely quantified, a heterogeneous group is obtained for which one can predict reading failure with a high degree of accuracy in nearly all cases (Thomson, 1975).

With regard to cerebral dominance, however, other researches have been concerned with its relationship to the wider issue of psychological function (Zangwill, 1960; Goldberg and Schiffman, 1972; Annett and Turner, 1974). Nevertheless, various aspects of research into cerebral dominance and information processing suggest a connection with reading ability. Research in information processing has produced obvious implications for reading research (Thomas, 1969; Fritzen, 1972; Haber, 1973). The relationships to be found between poor reading and problems of short-term memory have been investigated superficially (Rose, 1958; Alwitt, 1963; Birch and Belmont, 1965). Also Sperry (1968) thought it necessary for cerebral dominance to be established for information processing to proceed successfully in a sequential manner from the right to the left cerebral hemisphere. If Sperry is correct, the reading retardation associated with cross-laterality would seem, therefore, to be associated with impaired sequential information processing. Other recent advances in research on short-term memory have supported, and added to, Sperry's work and have important implications for research into problems of reading (Klatzky and Atkinson, 1971; Siegel and Allik, 1973; Allik and Siegel, 1974; Watkins and Wheeler, 1976).

In order explicitly to integrate and substantiate these implications it was, therefore, proposed that an experiment should be undertaken involving all the relevant factors—laterality, various forms of information, and an information-processing task requiring immediate storage or the use of short-term memory. Children were used who had passed the age of the developmental establishment of cerebral dominance which is about 7 years (Lenneberg, 1964 ; Goldberg and Schiffman, 1972). The children were also selected within an IQ range 110-120 because at that level children of the age range chosen should have mastered the basic processes of reading (Hage and Stroud, 1959). It was predicted that ill-established cerebral dominance, as indicated by problems of cross-laterality, would be related both to a limitation in information processing *irrespective of the type of information* (probably manifesting itself even with relatively small amounts of information) and significantly to reading retardation.

METHOD

Sample.

Two groups were obtained on the basis of laterality from a sample of 168 children from the western sector of Sheffield. The experimental cross-lateral group consisted of 10 children between the ages 8 years 5 months and 9 years 2 months (mean age=8 years 9 months) and non-verbal IQ between 110-120 (mean IQ=114·6) measured on the Raven's coloured progressive matrices. The children had been diagnosed as cross-lateral on the basis of the Harris test of lateral dominance (eyedness, handedness and footedness) together with ear dominance tests; and had obtained a score of zero scored on Thomson's criteria (Thomson, 1975). A control group of 10 children between the ages of 8 years 4 months and 9 years 2 months (mean age=8 years 10 months) and non-verbal IQ between 110-120 (mean IQ=115·5) was selected on the basis of consistent laterality. The two groups were not significantly different for age (T=·56; df=18) or performance on the non-verbal intelligence test (T=·59; df=18). All subjects were tested for reading age on the Schonell Graded Word Reading Test. Any child with hearing or sight problems was excluded from the study.

Apparatus.

Stimuli were presented in an Electronics Development two field card tachistoscope. The cards were presented at a distance of 490 mm. from the subject's eyes giving a subjective illumination at the eye of about 5 lux. A Behaviour Systems International Audio Generator Model 258 supplied a 4 khz supporting tone triggered by the start pulse generator and timer on the tachistoscope.

Stimuli.

The stimuli consisted of 180 cards in 9 sets of 20 cards as follows: 3, 5 and 7 digits; 3, 5 and 7 letters; and 3, 4 and 5 symbols (square, triangle, cross, diamond and circle). The sequence of all stimuli was ordered from random number tables. The stimuli were all 10 mm. high (Letraset didot 36 pt. Helvetica Medium).

Experimental Design.

A 2 (laterality) × 3 (form) × 3 (number) factor design, with repeated measures on second and third factors was employed. Each subject in both groups was presented tachistoscopically with all three categories of stimulus for the three different values of units. The order of presentation of the categories and values was randomised for each child. To avoid a differential fatigue effect, even though the order had been randomised, only one stimulus set was presented to a subject without rest. The average minimum time in m.sec for the subject to make a 100 per cent accurate verbal recall was obtained by the method of converging limits. A set of practice trials was given.

Procedure.

Children were tested individually and instructed to focus on the central black spot in the secondary field of the tachistoscope. The stimulus was presented centrally in the primary field. Their attentiveness was supported by a tone of 4 khz given 200 m.sec prior to the presentation of the stimulus. They were instructed to make an immediate verbal response after the termination of the stimulus presentation. They were told to guess if they were not sure exactly what stimulus had been presented.

RESULTS

The results and statistical analysis are presented in three sections—mean results and comparison between the two groups; an analysis of variance for the three factors of laterality, form and number; and finally, a consideration of some results about the reading ability of the two groups.

TABLE 1

TABLE OF MEAN EXPOSURE TIMES IN MSEC FOR CROSS-LATERAL AND UNI-LATERAL GROUPS, WITH THREE DIFFERENT FORMS, AND THREE DIFFERENT NUMBERS OF UNITS.

Conditions	Cross-Lateral (Time M Sec)		Uni-Lateral (Time M Sec)		Mann-Whitney ‘ U ’ Test (Difference between groups)		
	Mean	SE	Mean	SE	‘ U ’	nl/n2	P<
A digits 3	10·5	4·67	5·2	2·76	34	10,10	NS
5	885·5	147·38	339	94·18	12·5	10,10	·01
7	7825	1145·07	3210	475·44	12	10,10	·01
B Letters 3	58	22·82	3·0	·21	6	10,10	·001
5	2660	489·66	442	64·86	1	10,10	·001
7	10050	715·7	4830	713·37	7	10,10	·001
C Symbols 3	685·7	241·47	427·8	343·83	22	10,10	·025
4	3508	969·67	1381·0	374·12	26	10,10	·05
5	6860	978·57	3680·0	959·46	23	10,10	·025

A 2 (laterality) × 3 (form) × 2 (number of units: 3 and 5 units only) factor analysis of variance with repeated measures on second and third factors yielded the following results:

(1) Laterality was a significant factor in the performance of the groups (F=14·95; df=1, 18; P<·005).

(2) The type of information presented was a significant factor for both groups (F=43·70; df=2, 36; P<·001).

(3) The number of units of information presented was a significant factor for every type of information presented (F=72·14; df=1, 18; P<·001).

(4) A second order interaction factor of type or form of information and number was significant (F=24·54; df=2, 36; P<·001).

(5) A second order interaction factor of laterality and number of units was also significant (F=13·90; df=1, 18; P<·005).

(6) There were no other significant interaction factors.

TABLE 2

TABLE OF MEAN CHRONOLOGICAL AND MEAN READING AGES FOR UNILATERAL AND CROSS-LATERAL GROUPS.

Group	Chronological Age and Standard Error (in yrs., mths.)	Reading Age and Standard Error (in yrs., mths.)	Difference (in months)	Mann Whitney ' U ' test (between values)		
				' U '=	nl/n2=	P<=
Uni-lateral....	8 yrs. 9·7 mths. (±1·07 mths.)	11 yrs. 0·6 mths. (±4·62 mths.)	25·7 (±4·76)	9	10/10	·001
Cross-lateral ..	8 yrs. 8·8 mths. (±1·08 mths.)	9 yrs. 0·7 mths. (±9·17 mths.)	3·9 (±9·54)	41·5	10/10	NS
Difference between uni-lateral and cross-lateral groups			23·9	22	10/10	·05

DISCUSSION

Considering Table 1, a number of points are evident. The cross-lateral group take considerably longer than the uni-lateral group to process information correctly. This difference becomes increasingly marked as one increases the number of units of information and also the information set size of the types of unit, i.e., from numbers (10 alternatives), to letters (26 alternatives), or symbols (theoretically limitless), although it can be argued that the massive difference between the processing times for 3 and 5 symbols for both groups may be because the material is not easily verbally coded.

There is, however, an exception to this finding, namely for the processing time of 3 digits. This result is consistent with the second order interaction factor of laterality and number of units because, for small numbers of this type of unit, the different information processing abilities of the two groups do not manifest themselves until there is an increase in the load of information for the cross-lateral group.

The positive analysis of variance results for three factors of laterality, type of information presented, and number of units were obviously consistent with the hypotheses in the Introduction. The second order interaction factor information type and number of units was not surprising as the two combined give the total amount of information presented.

Considering Table 2, it follows that the uni-lateral group's reading age was significantly higher than their chronological age, considering their level of intelligence (Hage and Stroud, 1959). Equally so the results for the cross-laterals were consistent with previous research (Birch and Belmont, 1965; Allik and Siegel, 1974; Harris, 1957; Thomson, 1975; Gates and Bond, 1936).

What was surprising was that two groups balanced in all respects other than laterality, showed such a massive absolute and relative disparity in reading age (nearly 24 months) considering that the two groups were both well under 9 years of age.

These results when compared with other recent ones can help to clarify the relationship between cerebral dominance, laterality, short-term memory, information processing and reading ability. A number of researches have shown that information decoding is processed in a parallel manner in both cerebral hemispheres (Klatzky and Atkinson, 1971). However, information encoding is serially processed which requires information to be processed sequentially from right to left hemispheres. If cerebral dominance is ill-established this sequential processing is impaired and manifests itself as a limitation in information processing. This suggests either an under-utilisation of short-term memory or a limitation in the capacity of the short-term memory store.

Haber's (1973) model of information processing has as its central processing mechanism short-term memory as a common store for both visual and auditory information with a link to long-term memory and response production. It has been demonstrated in this experiment that there is a failing in short-term memory which is associated with ill-established cerebral dominance characterised by problems of cross-laterality. This explains Birch and Belmont's (1965) finding that poor auditory/visual integration is related to failure in reading.

This limitation in short-term memory store affects any kind of information processing task, not just reading. It presents a model to explain why children with ill-established cerebral dominance have difficulties with reading; and also why they are nearly always associated with difficulty in spelling and writing as they are all serially processed and require a relatively large store and fast process time in short-term memory.

REFERENCES

ALLIK, J. P., and SIEGEL, A. W. (1974). Facilitation of sequential short-term memory with pictorial stimuli. *J. exp. Psychol.*, 103, 567-573.

ALWITT, L. F. (1963). Decay of immediate memory for visually presented digits among non-readers and readers. *J. educ. Psychol.*, 54, 144-148.

ANNETT, M., and TURNER, A. (1974). Laterality and the growth of intellectual abilities. *Br. J. educ. Psychol.*, 44, 37-45.

BANNATYNE, A. (1971). *Language, Reading and Learning Disabilities.* Illinois: C. C. Thomas.

BELMONT, L., and BIRCH, H. G. (1965). Lateral dominance, lateral awareness and reading disability. *Child Dev.*, 36, 57-69.

BIRCH, H. G., and BELMONT, L. (1965). Auditory-visual integration, intelligence and reading ability in school children. *Percept. Mot. Skills*, 20, 295-305.

FRITZEN, J. (1972). Memory search and the recall of digit series. *J. exp. Psychol.*, 93, 410-416.

GATES, A. I., and BOND, G. L. (1936). Relationship of handedness, eye sighting, and acuity dominance to reading. *J. educ. Psychol.*, 27, 450-456.

GOLDBERG, H. K., and SCHIFFMAN, G. B. (1972). The role of dominance. In GOLDBERG, H. K., and SCHIFFMAN, G. B. (Eds.), *Dyslexia: Problems of Reading Disabilities.* New York: Grune and Stratton.

HABER, R. N. (1973). Information processing. In CARTERETTE, E. C. (Ed.), *Handbook of Perception* Vol. 1. New York: Academic Press.

HABER, R. N., and HERSHENSON, M. (1973). *The Psychology of Visual Perception.* New York: Holt, Rinehart and Winston Inc.

HAGE, D. S., and STROUD, J. B. (1959). Reading proficiency and intelligence scores, verbal and non-verbal. *J. educ. Res.*, 52, 258-262.

HARRIS, A. J. (1957). Lateral dominance, directional confusion, and reading disability. *J. Psychol.*, 44, 283-294.

HILLMAN, H. H. (1956). The effect of laterality upon reading ability. *Durham Research Rev.*, 7, 86-96.

KLATZKY, R. L., and ATKINSON, R. C. (1971). Specialization of the cerebral hemispheres in scanning for information in short-term memory. *Percept. Psychophys.*, 10, 335-338.

LENNEBERG, E. H. (1964). A biological perspective of language. In LENNEBERG, E. H. (Ed.), *New Directions in the Study of Language.* Cambridge, Mass.: MIT Press.

ROSE, F. C. (1958). The occurrence of short auditory memory span among school children referred for diagnosis of reading difficulties. *J. educ. Res.*, 51, 459-464.

SIEGEL, A. W., and ALLIK, J. P. (1973). A developmental study of visual and auditory short-term memory. *J. verb. Learn. verb. Behav.*, 12, 409-418.

SMITH, L. C. (1950). A study of laterality characteristics of retarded readers and reading achievers. *J. exp. Educ.*, 18, 321-325.

SPERRY, R. W. (1968). Cerebral dominance in perception. In ACS (Eds.), *Early Experiences in Childhood.* Washington: ACS Press.

THOMAS, H. B. G. (1969). Age and preferred information rates in a randomization task. *J. Child. Psychol., Psych.*, 10, 177-193.

THOMSON, M. E. (1975). Laterality and reading attainment. *Br. J. educ. Psychol.*, 45, 317-322.

WATKINS, E. J., and WHEELER, T. J. (1976). Laterality and short-term information processing: implications for reading failure. *IRCS (Res. Psychol.)*, 4, 139.

ZANGWILL, O. L. (1960). *Cerebral Dominance and its Relationship to Psychological Function.* Edinburgh: Oliver and Boyd.

ZANGWILL, O. L. (1962). Dyslexia in relation to cerebral dominance. In MONEY, J. (Ed.), *Reading Disability.* Baltimore: John Hopkins Press.

(*Manuscript received 16th August, 1976*)

Br. J. educ. Psychol., 47, 132-137, 1977

ICONIC MEMORY AND READING PERFORMANCE IN NINE-YEAR-OLD CHILDREN

BY R. J. RIDING AND J. C. PUGH
(*Department of Educational Psychology, University of Birmingham*)

SUMMARY. 36 9-year-old children were given a test of image persistence in visual sensory (iconic) memory, and the Neale Analysis of Reading Ability. In the iconic memory test the subjects viewed a white disc in a tachistoscope and were required to state whether or not the disc disappeared for a short interval which ranged from 10 ms to 800 ms. The shortest disappearance perceived was taken as a measure of icon persistence. The reading test gave scores for fluency, accuracy and comprehension. All three measures of reading performance were found to be significantly related to icon persistence. Short and long image persistence resulted in a reading age on the accuracy score that was on average 1·75 years below that for moderate persistence.

INTRODUCTION

THE reading process may be divided into three stages: (1) the registration in visual sensory memory of the images of the words scanned, (2) the analysis in short-term memory of the meaning of the words and sentences, and (3) the storage of the sense of the sentences in long-term memory. Such a division corresponds to current models of the learning process. In this paper the contribution of sensory memory to reading performance will be considered.

While the most obvious indication of a visual sensory memory is the visual after-effect in which the image of a bright light persists after it has been switched off, there is evidence that all stimuli are retained in image form for a short interval after their presentation. Treisman *et al.* (1975) surveyed the research supporting the existence of a short-lived sensory memory. The evidence may be summarised in three categories: (*a*) However short an observable visual stimulus, it appears to subjects to last for at least 120 milliseconds; (*b*) When two different stimuli are presented within 300 milliseconds of one another they tend to be integrated and seen as one. Similarly, the same stimulus presented intermittently at least three times per second appears to be continuous; (*c*) Information can be ' read off ' an image of a stimulus for several seconds after its presentation. The availability depending on the illumination and contrast of the stimulus. In a comprehensive study of iconic memory, Sakitt (1976) concluded that the storage of the visual image, or icon, is very largely located in the photoreceptors of the retina. The function of iconic memory appears to be to retain the sensory images in a fairly literal form long enough for the information they contain to be extracted and transferred to the next stage of processing.

Sperling (1960, 1963) found that images could be lost from iconic memory both by decay with time, and by being masked by the presentation of a second interfering stimulus. He observed that the icon retained a substantial amount of information for at least half a second after the presentation of a stimulus, although its clarity depended on the relative intensity of the stimulus. Further studies of loss by masking, reviewed by Kahneman (1968), showed that images could be lost by being masked by others presented either immediately before or after the test stimulus. In forward masking the presence in iconic memory of a previously presented image interferes with the next one presented by making its registration less clear, while in backward masking an existing image tends to be obliterated by the presentation of a further stimulus. For masking to occur the stimuli must be presented within a few hundred milliseconds of one another. Kahneman concluded that forward masking generally produced the greatest loss, although interference depended on the degree to which one pattern was superimposed on a second.

The results of studies of loss from iconic memory may be used to predict reading performance. In efficient reading the image of a word in iconic memory must be retained long enough for its features to be transmitted to the meaning analysis stage. Consequently, reading will be impaired if either the icon fades too quickly or persists too long. If the icon decays too quickly the reader's eyes will have to remain focused on each word until it is analysed, causing scanning to be jerky and slow. If the icon remains strong for a long interval it may forward mask the images of the words that follow. The most efficient condition for reading will be when the icon fades at a rate appropriate to the optimum speed at which the meaning of words can be analysed. If individuals differ in icon duration then it is likely that this variation will be reflected in reading performance.

No studies of the distribution of individual variation in iconic memory ability have been found. Iconic performance in normal children has been compared with that in dyslexics (Stanley and Hall, 1973), and with adolescent retardates (Welsandt and Meyer, 1974), but neither study found a difference in icon retention. Guthrie (1972) found a significant relationship between reading ability and short-term visual memory in children. Visual memory was tested by presenting a series of geometrical forms followed by immediate recall or reconstruction of the series. While this study did not investigate iconic memory directly, any effect it had was included in the result. An experiment is required to determine whether icon persistence varies between individuals and how it affects reading performance.

In the study of iconic memory performance in children the effect of age must be controlled for. Several investigations have shown that the speed of decay of icons in memory increases with age. For instance, Pollack *et al.* (1969) studied the threshold for the detection of a dark interval between two flashes of light in children whose ages ranged from 6 to 17 years. They found that the oldest subjects could detect an interval that was 30 per cent shorter than that perceived by the 6-year-olds, and concluded that the difference in performance was due to a decrease in the persistence of the image in iconic memory caused by physiological ageing. Using a masking method, Liss and Haith (1970) found that 5- and 10-year-old children required a longer interval between a test stimulus and a mask to say correctly whether the test stimulus contained a vertical or a horizontal bar, than did adults. In both the masking conditions, adults responded correctly with an interval between stimuli that was much lower than that required by the children. Gummerman and Gray (1972) showed that while icon persistence decreases with age, overall information processing efficiency improves as children get older and this compensates for the deterioration in iconic memory.

For use with children a method of measuring iconic memory performance should be within the child's capability and validly test only that ability. Bongartz and Scheerer (1976) have warned that in the partial-report method and in methods using a stimulus array, short-term visual memory is likely to be implicated in addition to iconic memory. A suitable test would appear to be the dark interval threshold method of the type employed by Pollack *et al.* in which the subject only has to report whether there is a dark interval or not. The minimum dark interval perceived reflects the duration of the icon.

The aim of the present study was to investigate individual differences in the icon persistence of children and compare them with their reading performance.

METHOD

Sample.

The sample consisted of 36 children (18 boys and 18 girls) whose age was within two months of 9 years 7 months. They were drawn randomly from the pupils of that age range in an urban middle school. The regression equation given by Pollack *et al.* relating age and the dark interval threshold gave ·85 milliseconds as the decrease in the threshold when age is increased by four months.

Materials and Procedure.

Each child was given individually the test of iconic memory followed after four days by the test of reading performance.

Iconic Persistence Test. A modified version of the dark interval threshold test employed by Pollack *et al.* was used. Each child sat in front of a Colne two-field tachistoscope (model CET/2) and viewed the stimulus which was a 1 cm diameter white disc on a black background. The disc was continuously presented except when the experimenter switched for a very short interval to the second field which was completely black. The subject viewed the illuminated disc and was instructed to say whether or not it disappeared and came back on again.

After doing some practice items during which the experimenter ensured that the child fully understood what he should do, there were 18 trials in which the experimenter presented dark intervals ranging from 10 milliseconds to 800 milliseconds, in random order. When the field was changed a 'flicker' was apparent and this was used to indicate to the subject when a trial occurred. The subject looked into the tachistoscope for the duration of the test to ensure that accommodation to the illumination was kept at a steady state. After each trial the experimenter asked whether or not the disc had disappeared when the flicker occurred. The dark intervals used (in milliseconds) were 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 115, 130, 150, 200, 300, 400, 600 and 800. A pilot investigation had shown that no subject of this age reported seeing an interval of 20 milliseconds or less, while all observed intervals of 300 milliseconds and more. For each subject the smallest dark interval that could be distinguished was taken to be an indication of icon persistence. Obviously, some trace of the icon remained for longer than the minimum interval seen but it was not sufficiently strong to prevent the interval being perceived. In most cases, as the dark interval decreased, there was a clear change from observing the interval to not seeing it. When there was a period of uncertainty and a particular interval was not observed, but a smaller one was, the icon persistence was taken to be the mean value of the range of uncertainty. The mean range of uncertainty was small: 5·97 milliseconds (SD 11·39).

Reading Ability Test. The Neale Analysis of Reading Ability Form A (Neale, 1966) was used. This gave scores for reading fluency, accuracy and comprehension. The raw scores were converted into reading quotients, using the norms provided, to control for the small age differences of the subjects.

RESULTS AND DISCUSSION

Considerable individual differences in icon persistence were observed. These ranged from 20 milliseconds to 300 milliseconds with a mean of 102·97 milliseconds (SD 55·64). Since no significant sex differences in either icon persistence or reading ability were found, in order to compare iconic memory with reading performance the subjects were divided into three groups according to their icon persistence as indicated by the shortest dark interval perceived. The range of persistence and the number of subjects for each group was: short 30—70 milliseconds (mean 55·42 milliseconds, SD 13·22, N=12); moderate 71—100 milliseconds (mean 87·73 milliseconds, SD 10·57, N=11); and long 101—300 milliseconds (mean 159·77 milliseconds, SD 52·99, N=13). The mean quotients for the three measures of reading ability for the short, moderate and long icon persistence groups are given in Table 1.

TABLE 1

THE EFFECT OF ICON PERSISTENCE ON READING PERFORMANCE.

Reading Measure	Mean Reading Quotient of Icon Persistence Group					
	Short		Moderate		Long	
	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)
Fluency	104·6	(18·36)	124·6	(14·53)	110·2	(21·32)
Accuracy	90·4	(7·15)	110·6	(12·31)	94·3	(8·50)
Comprehension........	91·3	(13·62)	110·6	(13·79)	92·8	(9·25)

It was found that moderate icon retention was related to much better reading performance than either short or long persistence on all three measures, and that long persistence gave marginally better reading scores than short. A one-way analysis of variance was performed for each measure and these are summarised in Table 2.

TABLE 2

ANALYSES OF VARIANCE OF READING MEASURES.

Reading Measure	F	df	P<
Fluency	3·57	2, 33	·05
Accuracy	14·56	2, 33	·001
Comprehension	8·78	2, 33	·001

Icon duration was significantly related to all three aspects of reading. It seems likely that the best reading performance is obtained when the icon persistence is moderate because the image of each word remains long enough to be analysed but not so long that it interferes with the reception of subsequent words. When the persistence of the image is very short the reader will probably be hindered by having to focus on each word until its meaning has been determined, resulting in jerky scanning. With icons that are too long lasting some masking is likely to occur when the images of newly received words suffer interference from previously registered ones.

Considering the first of the measures of reading performance, *fluency* or speed of reading, if the image decays very quickly the reading will have to be in terms of one word, or a very small group of words, at a time and will be inefficient. In the case of long icon persistence the reader will have to scan slowly to allow sufficient time for the image of a word to decay before moving to the next if masking is to be avoided. Both of these conditions will affect the fluency of reading. The effect was, however, less reliable than for the other measures perhaps because as an indication of reading ability speed does not distinguish between those children who are both quick and accurate, and those making many mistakes. Further, since the subjects were not instructed to read as quickly as possible there is no certainty that all made an equal effort to do so. The *accuracy* of reading was found to be very reliably related to icon persistence. It is suggested that this was because the fast fade of the icons in the short persistence subjects caused the words to have been misread, particularly if the subject tried to maintain a reasonable reading speed, and in the long persistence children forward masking would have caused the 'misreading' of some words from iconic memory when the presence of two images superimposed on one another made them indistinct. *Comprehension* performance was also related to image persistence but less significantly than accuracy, presumably because the quality of comprehension depends not only on the information that is received via iconic memory, but also on the efficiency of the higher meaning analysis processes and the contents of long-term memory.

The size of the reading performance difference between the moderate image duration group and those of short and long persistence was large enough to be of practical importance. The mean reading ages of the 9-year-old subjects on the accuracy test were: 10·59 years for the moderate group, but only 8·66 years and 9·03 years for the short and long retention groups, respectively. Follow-up analyses of variance showed that both the short and the long group means differed significantly from the moderate group's performance (respectively, $F=21{\cdot}29$; df 1, 21; $P<{\cdot}001$; $F=14{\cdot}51$; df1, 22; $P<{\cdot}01$) although the difference between them was not reliable ($F=1{\cdot}52$; df 1, 23; $P>{\cdot}25$).

Further research is necessary to determine the precise form of the relationship between icon duration and reading performance, and how it interacts with the age of the subjects. The interpretation of the results presented here could be checked by comparing reading performance and persistence at various externally controlled rates of presentation and exposure times of words. For the short icon persistence subjects, increasing the exposure time of the words for a constant presentation rate should improve reading accuracy while decreasing exposure time when presentation is at a moderate rate should improve performance for the long duration group. If such a subsequent study confirmed the present explanation it might be worth experimenting with different print-background contrasts to make it easier for short and long duration pupils to read. Presumably a very high contrast between type and background would help the short duration children by slightly increasing the relative persistence of the icons, while low contrast print, say dark grey on light grey, may help the long retention subjects by reducing icon duration.

REFERENCES

BONGARTZ, W., and SCHEERER, E. (1976). Two visual stores and two processing operations in tachistoscopic partial report. *Q. J. exp. Psychol.*, 28, 203-219.

GUMMERMAN, K., and GRAY, C. R. (1972). Iconic storage and visual information processing. *J. exp. Child Psychol.*, 13, 165-170.

GUTHRIE, J. T. (1972). Visual sequential memory in reading disability. *J. Learn. Dis.*, 5, 41-48.

KAHNEMAN, D. (1968). Method, findings and theory in studies of visual masking. *Psychol. Bull.*, 70, 404-425.

LISS, P. H., and HAITH, M. M. (1970). The speed of visual processing in children and adults: effects of backward and forward masking. *Percept. Psychophys.*, 8, 396-398.

NEALE, M. D. (1966). *The Neale Analysis of Reading Ability.* London: Macmillan.

POLLACK, R. H., PTASHNE, R. I., and CARTER, D. J. (1969). The effects of age and intelligence on the dark-interval threshold. *Percept. Psychophys.*, 6, 50-52.

SAKITT, B. (1976). Iconic memory. *Psychol. Rev.*, 83, 257-276.

SPERLING, G. (1960). The information available in brief visual presentations. *Psychol. Monogr.*, 74, 11 (Whole No. 498).

SPERLING, G. (1963). A model for visual memory tasks. *Hum. Fact.*, 5, 19-31.

STANLEY, G., and HALL, R. (1973). A comparison of dyslexics and normals in recalling letter arrays after brief presentations. *Br. J. educ. Psychol.*, 43, 301-304.

TREISMAN, A., RUSSELL, R., and GREEN, J. (1975). Brief visual storage of shape and movement. In RABBITT, P. M. A., and DORNIC, S. (Eds.), *Attention and Performance V.* London: Academic Press.

WELSANDT, R. F., and MEYER, P. A. (1974). Visual masking, mental age and retardation. *J. exp. Child Psychol.*, 18, 512-519.

(*Manuscript received 3rd November*, 1976)

Br. J. educ. Psychol., 47, 138-148, 1977

THE READING ABILITY OF DEAF SCHOOL-LEAVERS

BY R. CONRAD

(*Medical Research Council Applied Psychology Unit, Cambridge*)

SUMMARY. Almost every child aged 15-16½ years being educated in schools for the deaf and partially hearing in England and Wales, was given Brimer's Wide-Span Reading Test which assesses prose comprehension. After excluding certain categories of children, results were available for 355 children. With degree of hearing loss as the main independent variable, results are presented in terms of mean and median reading age, and in reading age percentile form. Reading is significantly affected by degree of deafness and by non-verbal intelligence. Though the presented reading ages are with reference to hearing norms, other criteria for assessing the reading performance of deaf children are discussed.

INTRODUCTION

ALTHOUGH there have been few empirical studies published relating to the reading ability of deaf children, it is widely assumed by teachers of the deaf that the prose comprehension of their pupils remains poor. Vestberg Rasmussen (1973), referring to Danish children, says frankly: " We are forced to admit that we cannot teach these deaf children to read " (p. 32). Watson (1967) expresses ' grave concern ' and ' perplexity ' because: " . . . many deaf children become retarded in reading " (p. 134), a phrase which in the light of results to be presented may even seem euphemistic. This concern is widespread, and Myklebust (1964) refers to a 7–8 year retardation at age 15 years. Yet comprehensive description of the relationship of degree of hearing loss to reading is sparse generally and lacking in Britain. In other words, though it is believed that deaf children in Britain are very retarded readers, that statement about comprises the extent of knowledge except for pockets of small special samples.

One might have thought from this, that problems concerned with teaching reading to deaf children would have long been a vigorous research area. Not at all. The *Deafness, Speech and Hearing Abstracts*, begun in 1960, did not justify an indexed subsection on reading until 1973. A seminal and selective international congress on educational methods for the deaf (Ewing, 1960) heard not a single discussion on reading. Ten years later, the *Proceedings of the International Congress on the Education of the Deaf* (1970) again yield no direct evidence concerning reading. The enquiry set up in Britain on the possible role of signing in the education of the deaf (Lewis, 1968) has no more to say on reading than that everyone attaches importance to teaching it. The main aim of this article is to describe quantitatively some aspects of the state of affairs in England and Wales.

There have, of course, been a number of similar studies, all—as is the present study—limited in some degree. Comprehensiveness of sample is an obvious and costly problem. Murphy (1957), in a very thorough study of the complete population of 12-year-olds in schools for the deaf in England, only used a word recognition test to assess reading. He also reported that the test correlated highly with a comprehension test using a hearing sample. But there is no way from this to derive the comprehension ability of the deaf population. Hamp (1972), in another large-scale enquiry, did use a sentence completion test amongst others. But he spread his 367 subjects throughout the 9–16-year age range. Thus, he has insufficient data to separate out independently the crucial variables of age and hearing loss.

More commonly, studies of reading by the deaf have been small in scope and of minor evaluative interest. Owrid (1970) included a word recognition test in a study of 27 children aged 5–11, and gave no audiological information about them. Hine (1970) presents reading comprehension values for 100 selected partially hearing children aged 8–16. But the reading age of the older children, which is relevant to this study (about 11 years), is representative only of one school. Redgate (1972), in his report on the use of i.t.a. with young deaf children, quotes some unpublished data showing the mean reading age of a small sample of 15-year-old deaf children to be about 8 years. But no other information is provided.

Vernon (1976) summarises a number of American studies which are concerned with reading ability in relation to early use of signing. Appropriately selected subject samples are used and the respective quoted reading ages presented are of limited value in a general sense. So far as the USA is concerned, the definitive study now is a nationwide description of educational attainments aiming at a 100 per cent sample of deaf school children. DiFrancesca (1972) presents data on reading comprehension (and other) ability for some 17,000 subjects, with breakdowns by age, and hearing loss. Jensema (1975) elaborates on these data in a further breakdown by cause of deafness, sex, additional handicaps and ethnic background. Two weaknesses may be thought to mar this very important research. First, the test scores were supplied by the schools. Second, reading comprehension levels were assessed by use of the Paragraph Meaning subtest of the Stanford Achievement Test. Because of the wide range both of chronological-age and reading-age of the subjects, five different levels of the subtest are involved. For comparative purposes, these then have to be brought together into a necessarily arbitrary unitary weighted score. Nevertheless, the outcome cannot be far from being a very thorough and accurate description without peer. The DiFrancesca study provides invaluable comparative data against which other educational programmes can be evaluated.

The results on reading to be presented here are not intended to emulate DiFrancesca's; nor are they to be regarded as a UK replication. They are taken from a wider study concerned with an examination of a number of verbal skills of deaf children and the assessment of reading ability was linked to broader issues. But the study was also designed to provide descriptive data on levels of attainment at the end of formal school education. The population tested comprises almost all children receiving special education because of hearing impairment and amenable to formal education, and who were aged 15–16½ years at the time of testing. The sample, therefore, is limited only by age: the age by which statutory education has ended its role. The data should be regarded as an evaluation of the reading ability of hearing impaired school-leavers, and because of the comprehensiveness of the sample, they may also be regarded as an evaluation of the educational treatment provided for these children. The children tested were all in special schools for the deaf or partially hearing. Children being educated in partially hearing units attached to ordinary schools were also tested and the results will be published at a later date.

METHOD

Schools.

In England and Wales, there are 41 special schools for the deaf or partially hearing catering for secondary-aged children, and where conventional formal teaching is considered to be appropriate. Some are day schools, some residential and some mixed. Testing was carried out in 38 of these schools. One small residential school declined to co-operate, one Education Authority refused access to its school and a day school with three relevant children had to be omitted for administrative reasons. Testing was completed between June, 1974, and January, 1976.

Sample.

At every school, all children were seen whose birthday at the date of testing fell between such dates that they were aged 15–16½ years. They are, therefore, the school leavers. Children older than this, and still at school, were excluded to avoid selection biases confounded with age once the statutory school leaving age was passed. The tested sample was totally unselected and almost 100 per cent of the target population. However, some children have been excluded from the results which follow, viz: children known to have had normal hearing up to at least the age of 3 years (N=42); children from homes where the language normally spoken was not English (N=18), but children from ' signing ' homes were included; children for whom so little medical or audiological information was available that they would have been unclassifiable (N=16); children who were totally untestable (N=10); a few children absent from school or sick. The final sample totalled 355.

Reading Test.

Attention has been drawn to the problems which arise when standard narrow-range reading tests suitable for homogeneous hearing populations are used with deaf children with great variance in reading ability. That was one difficulty to overcome. Secondly, a good many reading tests of the sentence-completion type provide, say, four alternative words—one of which is correct. It may be extremely difficult in test design to provide four words of equal difficulty to children with impoverished vocabulary as deaf children are—particularly when, as so often, they are all of the same part of speech. It is easy for such a test to become one of no more than word recognition (Myklebust 1964). Thirdly, it was important to be sure that children were not already familiar with the test.

Brimer's Wide-Span Reading Test (1972) met the requirements. This is an 80-item test standardised for reading ages of 7–16 years in a single test. An item consists of one sentence with a word missing, and a second sentence which contains the correct word. There may, therefore, be many familiar words in the second sentence and the correct word may be employed in a part of speech different from that in the target sentence. A greater complexity of linguistic fluency is tested in a much simpler manner than in the majority of sentence-completion reading tests. For the present purpose, two minor modifications were made. Two items, one of which was reached by few children, were omitted because the answer happened to depend on phonological characteristics of words. The second change was in the practice items. The standard three practice items were preceded by three even easier items, and one word was changed in a practice item to evade what is for deaf children the particularly difficult discrimination between ' off ' and ' of.' The effect of these changes would be indiscernible in the results. Two scoring procedures are available according to whether the test is given with a time limit or not. In the present case no time limit was set, and a ' stopping-rule ' was operated which effectively guaranteed that every child worked far beyond his limit of comprehension. This limit is determined by rules set out in the test manual.

Procedure.

Although the children were sometimes seated in groups which never exceeded five, test instructions were administered individually either by a qualified teacher of the deaf or by a psychologist. The instructions were given in any manner which would aid understanding. In general, the tester worked directly with the child over the first two or three practice items and then left him to complete the practice set. These were then checked, errors discussed, further clarification given if necessary, and the child set to work unaided on item 1 of the test proper. No further help was given. Evidence that the instructions had been understood came largely from scrutiny of the practice items, and then only 12 children failed unaided to answer item 1 correctly. The results which follow are based only on children who demonstrably did understand the nature of the task.

Scoring.

The norms, in the form of a reading age starting at 7:0 (years: months) are based on the furthest point of the test at which a child is correct on three out of four consecutive items. Since there were always at least two invigilators present it was easy to detect, by operation of a rule, when it was pointless for a child to continue, bearing in mind the fact that the test has a ceiling reading age of 16:0.

Other information.

On some other occasion, but within a day or two, each child completed Raven's Standard Progressive Matrices (1960), again with individual instruction. Information concerning hearing loss and medical history was obtained by interview with the school head and from relevant records.

RESULTS

A summary of the principal results is shown in Table 1. Whilst the relationships of many variables to reading ability are of interest and educational significance, the main aim of this article is to describe that between reading and degree of deafness.

The hearing loss of each child was taken from the latest better-ear audiogram available at the school, and averaged over the frequencies: 250, 500, 1,000, 2,000, 4,000 Hz—giving an average hearing loss in dB (ISO). The five bands of hearing loss used in Table 1 are the outcome of extensive analysis of variance of the pattern of hearing loss at individual, and combinations of frequencies with respect to speech quality. The bands used represent the best predictors of speech quality from pure-tone audiograms (Kyle, 1977). Especially relevant is the 85dB point, which is often taken to represent the hearing loss at which a child is considered to be profoundly deaf (Quigley, 1968 ; Vernon, 1976).

Conventionally, reading ability is presented in the form of an arithmetic mean reading age, or mean reading grade, based on norms for hearing children. Unless reading ability is normally distributed in the sample, an average can be seriously misleading if it is used as a descriptor and an estimate based on percentile values is likely to be more useful. Table 1 illustrates this for the present data by giving both mean and median values of reading age in months. The discrepancy is particularly great for the deafer children. At 86–95dB, the mean reading age is about 9 years. But more than 50 per cent of children fail to attain a reading age of 8 years.

Kruskal-Wallis one-way analysis of variance indicates that reading ability depends significantly on degree of hearing loss ($\chi^2=53$, $P<\cdot001$). The results of Mann-Whitney tests show that the only differences which do not reach the ·05 significance level, are between the groups 0–65dB and 66–85dB and between 86–95dB and 96–105dB. There is no significant effect of hearing loss on non-verbal intelligence, the five groups having effectively the same mean IQ. This is of some importance, because the variances of the respective IQ scores are similar throughout the hearing loss range, and none significantly different from an expected normal distribution. This fact, together with the similarity of means, implies that there is not a concentration of low IQ children at any particular hearing loss level. Superficial observation in schools easily leads one to imagine that classes of partially hearing children are over-burdened with children of particularly low intelligence. These data do not support that.

The product-moment correlation between reading age and score on Raven's Matrices was calculated separately for each level of hearing loss. All the correlations are significant beyond the ·05 level. The apparent decrease with deafness in the size of the correlation is probably due to the increasing proportion of children unable to read at all.

TABLE 1

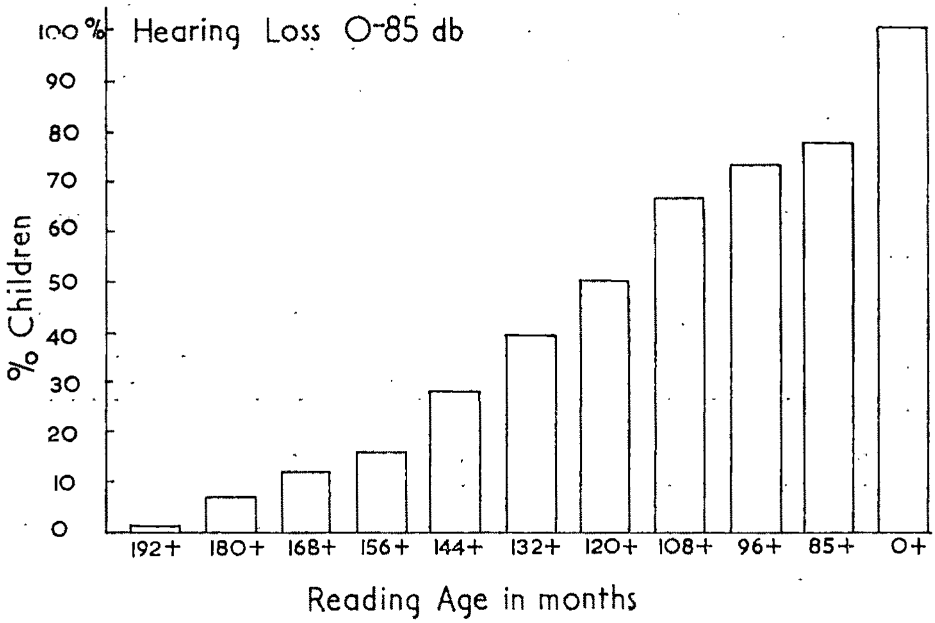
HEARING LOSS, READING AND INTELLIGENCE.

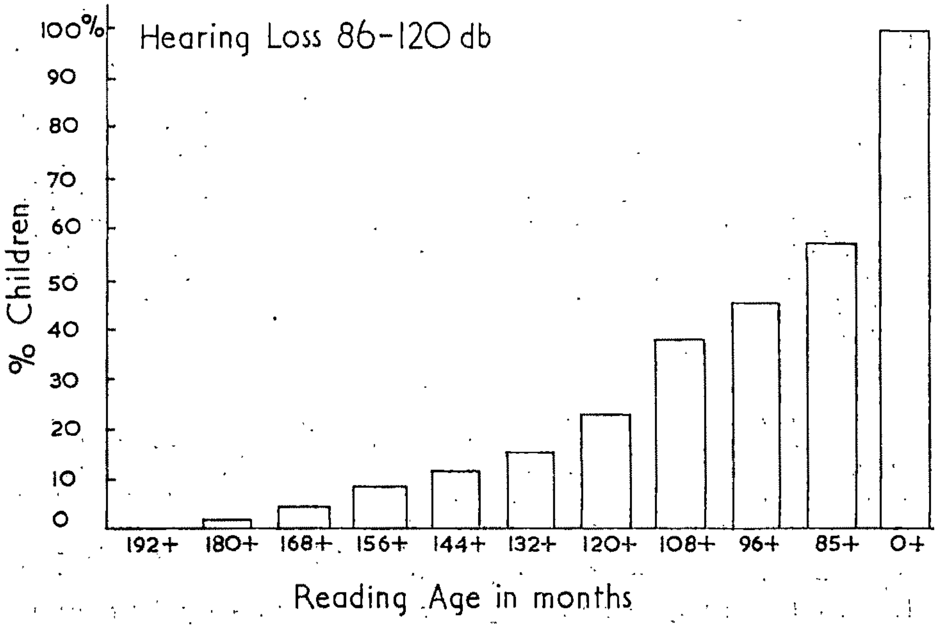
	Hearing Loss (dB) 0—65	66—85	86—95	96—105	106+	Overall
(N)	68	82	55	84	66	355
Mean IQ	96	99	98	98	94	97
Median Reading Age (Months).	128	112	94	92	84	108
Mean Reading Age (months)	124	121	109	107	99	112
(sd)	(32)	(34)	(32)	(28)	(23)	
Reading × Intelligence correlation	·53	·48	·45	·48	·30	

The impact of deafness on reading ability is illustrated with more force perhaps in Figure 1. Here, to preserve numbers, hearing loss groupings are collapsed to provide only two levels divided at 85dB. The histograms then show cumulatively the percentage number of children reaching different levels of reading age in months. It should be recalled that the Wide-Span Reading Test as used here is concerned only with prose comprehension, and has a floor of 84 months; the lowest detectable reading ability appears as a value greater than that. Figure 1 then shows that almost 50 per cent of the children with hearing loss greater than 85dB have no reading comprehension at all. This would be the case with a sample of hearing children only if they had a mean chronological age of 7 years. For the less deaf children the comparable value is about 25 per cent. Taking the entire population of children of school leaving age in schools for the deaf, it can be seen that some 40 per cent are totally illiterate for prose comprehension. At the other end of the range, in the partially hearing group, 8 per cent of children have a reading age commensurate with their chronological age. In the entire population, there are 4½ per cent of such children. The value for any other reading age can, of course, be read off.

FIGURE 1

CUMULATIVE FREQUENCIES OF READING AGE IN MONTHS OF CHILDREN AGED 15—16½ YEARS IN SCHOOLS FOR THE DEAF.





Since the correlations between reading and intelligence shown in Table 1 show a significant association, the reading ability of a deaf child evidently depends at least on his degree of deafness and his level of intelligence. There is no anomaly in the fact that the intelligence test is non-verbal. It has been well established (Raven, 1960) that the test correlates highly with verbal intelligence. But reading ability also depends significantly on hearing loss. At the very least one might assume that of all of the factors likely to contribute to reading ability at the end of schooling, these two must be given a high rating. Table 2 shows the outcome of considering these two factors together.

TABLE 2

PERCENTILE VALUES OF READING AGE IN MONTHS BY HEARING LOSS AND INTELLIGENCE.

Percentile	Hearing Loss			
	85dB or less		86dB or more	
	below average IQ	above average IQ	below average IQ	above average IQ
75th	132	164	100	142
50th	110	142	84	112
25th	84	100	84	86
(N)........	90	60	125	77

NOTE—Three subjects did not take the intelligence test.

Table 2 in effect takes the data of Figure 1 and divides it into below- and above-average intelligence; for this purpose a raw score of 45 or more on the Progressive Matrices is regarded as above average for a normal population. To illustrate the way that this table may be used, it shows for example, that of the population of children who are below average intelligence and with hearing loss less than 86dB, 25 per cent, but no more, reach a reading age of 11:0 years. The best performance is that of children with hearing loss less than 86dB and above average intelligence. Half of them reach a reading age of 11:10 years (142 months), and 25 per cent of them achieve a reading age of 13:8 years (164 months). However, the numbers of subjects involved should be noted. This group of relatively good readers totals 15 children in the entire school population of the age in England and Wales. At the other extreme, of the profoundly deaf children who are also below average intelligence (35 per cent of the population), only 25 per cent of them will achieve a reading age of 8:4 years; most of the other 75 per cent of this section of the deaf population will have in effect no prose comprehension at all.

DISCUSSION

The statement of reading ability of deaf children presented in Table 1 implies the use of the most commonly accepted comparative standard, but which is not necessarily of greatest value. It compares a number of deaf populations with an 'average' hearing child. For example, the least deaf group read about as well as the average hearing child aged 10½ years; the most deaf of the deaf groups have the same reading ability as the average hearing child aged 8:3 years. Remembering that the deaf children are all older than 15 years, then a very severe reading deficit, or rather reading difference, is obvious. Furth (1966) has pointed out the impropriety of regarding this difference as a reading deficiency rather than as an indication of linguistic incompetence. Though it seems hard to accept the mutual exclusivity of these two terms, the distinction does draw attention to the need to justify the use of this particular but universally used criterion.

If reading is considered as a tool for achieving something else, such as information or knowledge, then no particular virtue attaches to any specific level of reading ability until that level has been justified as necessary for the purpose. In other words if it is considered that adequate reading is represented by a reading age of 9:0 years, then 50 per cent of the present sample are adequate readers regardless of their apparent ' deficit.' But if the criterion is set at, e.g., 11:0 years, then one notes that 75 per cent of the children tested lack adequate reading. A second available comparative standard, therefore, would be one which relates to literacy rather than to the reading ability of hearing children.

Furth (1966) has suggested that " . . . a functionally useful ability to read. . . " is reached at the end of the USA 4th Grade, i.e., about 11:0 years (p. 461). Bullock (1975) summarises a number of viewpoints on this issue. He cites Moyle (1973) as regarding a reading age of 13 years to be that required to read the simplest daily newspaper. Later Bullock suggests that " . . . it has often been found that . . . the lowest grade of difficulty at which complex subject matter can be written approximates to a reading age of about 15 " (p. 11). More realistically, one might think, Bullock refers to a Ministry of Education pamphlet which defines an illiterate person as one with a reading age of less than 7 years, and a semi-literate a person with a reading age between 7 and 9 years. Which criterion of literacy one adopts is clearly a matter of opinion and particular purpose. The data presented permit an assessment to be made of the reading ability of deaf school-leavers on the basis of one's preferred criterion. Choosing one is not a goal of this paper. If Bullock's ' illiterate ' value of 7 years is accepted, then 35 per cent of this population are illiterate and another 20 per cent are no better than semi-literate. If Furth's more severe criterion is adopted, then 75 per cent of this population must be regarded as ' functionally ' illiterate. No more than 4 per cent of all the pre-lingually deafened children leaving special schools in England and Wales achieve the most testing criterion of all—the ability to understand ' complex subject matter,' and of the profoundly deaf, no more than 2½ per cent. All these figures should, of course, be contrasted with a figure of 50 per cent for hearing children. Finally, on this point, Hammermeister (1971) tested the reading ability of deaf subjects up to 13 years after leaving school and found no improvement at all on comprehension.

There is a third comparative standard which becomes useful at this point. Whether or not one regards the outcome of using the literacy criterion as being socially or educationally acceptable, it is still relevant to consider whether other educational systems achieve better performance. Ahlström (1971), referring to the reading ability of deaf Swedish children, uses phrases such as ' very low.' Nordén (1975) also reports that Swedish deaf school-leavers ' barely achieve ' a Grade 4 (i.e., about 10 years) reading age. Vestberg Rasmussen's (1973) pessimistic comment with respect to Danish deaf children was mentioned earlier. For each cross-cultural comparison reference to normative data is probably the most convenient.

The most comprehensive psychometric data come from the USA, and since the children are also taught to read English meaningful comparison is possible.

The classic data of Wrightstone *et al.* (1963) which were collected in 1959, have been superseded in usefulness by the DiFrancesca (1972) study on nearly 17,000 children of whom some 2,800 were aged either 15 or 16 years. Some problems arising from the use of the Stanford Achievement Test were referred to in the Introduction, but DiFrancesca does provide a weighted grade equivalent average score by age, though not by degree of hearing impairment. Combining the two ages 15 and 16 years yields a median reading grade of 3·15, which approximately would represent an age of 110 months. The overall median value shown in Table 1 here is 108 months. There is, therefore, a remarkable concordance. Of course, great caution is necessary when assessing its significance. But one might note that the distribution of hearing losses in the two populations is very similar. For example, the American population has 34 per cent of children with average hearing loss greater than 98dB, which is one of the cut-offs used. The population of this study has 35 per cent of such children. Furthermore, the average intelligence of the two populations is unlikely to be very different and will largely depend on the respective proportions of children unable to cope at all in ordinary education programmes, and therefore excluded from both studies. The principal factor likely to affect the validity of the comparison is the use of different reading tests. Since neither test is culture-free, this remains an insoluble problem the size of which readers must judge for themselves. On the criterion of comparability with other educational programmes therefore, the reading performance of deaf children in England and Wales is seen to be very similar to that of American children, and is the best *ad hoc* comparison likely to be possible. It is highly unlikely that the inclusion of children being educated in partially hearing units in this country will substantially affect this comparison after degree of hearing loss is taken into account. Whether reading performance in the two countries is similarly good or similarly bad has to be judged by criteria other than mere agreement; literacy has been suggested as one criterion. On any of the definitions discussed earlier, both populations are reading very poorly.

This suggests one more approach to the problem of assessment: given the nature of the handicap, what ceiling might be expected? This would have to depend on the outcome of a theoretical analysis of what is involved for a deaf child in learning to read the English language. On this basis a median reading age of about 9 years may, in fact, represent a theoretical limit. It has often been suggested for example that the poor phoneme-grapheme correspondence of English creates major problems for the deaf. But, in fact, there is no evidence that deaf children learning to read in other languages read better. Indeed, Moores (1972) has reported that in spite of the very close phoneme-grapheme correspondence in the Russian language, the Soviet Union found it necessary to abandon strictly ' oral ' methods of teaching reading to the deaf. At present few insights are current as to what kind of theoretical issues would merit study except in the broadest terms. Myklebust (1964) for instance refers to the need " . . . to foster association between the written and speechread forms of the word " (p. 275). Furth (1966) stresses the linguistic incompetence of the deaf. But Moores (1971) using a ' cloze ' procedure, states that 10-year-old hearing children use different language structures in reading from much older deaf children with a reading age of 10 years. The precise nature of the linguistic incompetence and its significance for reading remains obscure.

Still, inevitably, at a speculative level, Conrad (1972, 1973) has reported that whereas some deaf children memorise word lists using a speech code, many others do not. If the latter memorise words on the basis of what they look like, they would be handicapped by the limitations of sequential short-term visual memory (Paivio, 1971). Furthermore, such children would be particularly penalised by a phonic reading programme which would tend to be used in ' oral ' schools.

At first sight, and possibly at last sight also, deaf children leave school with what appears to be appallingly poor reading comprehension. But although it is an unambiguous criterion, perhaps the least useful one is the simple numerical discrepancy in comparison with hearing children. But because this disrepancy is so great it is unlikely that even the most radical changes in teaching methods will eliminate it. More realistic goals may be needed, and these have never been defined in a pedagogically useful way. Goals based on literacy offer some scope here, and then the goal can be varied from child to child according to his anticipated social needs. It is at that point that theoretical analysis of reading by the deaf will be crucial, since no goal can be achieved unless the learner-reader has available the appropriate cognitive organisation.

ACKNOWLEDGMENTS—Data collection was assisted by Dr. J. G. Kyle, Miss M. McKenzie, Miss A. Morris and Mrs. B. C. Weiskrantz. Dr. Kyle also advised on the statistical treatment of data. I am greatly indebted to the staffs of the 38 schools who provided facilities for testing.

REFERENCES

AHLSTRÖM, K. G. (1971). *On Evaluating the Effects of Schooling.* Proceedings of the International Congress on Education of the Deaf. Stockholm, 1970.

BRIMER, A. (1972). *Wide-Span Reading Test.* London : Nelson.

BULLOCK, A. (1975). *A Language for Life.* London : HMSO.

CONRAD, R. (1972). Short-term memory in the deaf: a test for speech coding. *Br. J. Psychol.*, 63, 173-180.

CONRAD, R. (1973). Some correlates of speech coding in short-term memory of the deaf. *J. Speech Hear. Res.*, 16, 375-384.

DIFRANCESCA, S. (1972). *Academic Achievement Test Results of a National Testing Programme for Hearing-impaired Students. United States: Spring,* 1971 *(Report No.* 9, *Series D)* Washington, DC: Gallaudet College, Office of Demographic Studies.

EWING, A. W. G. (Ed.) (1960). *The Modern Education Treatment of Deafness.* Manchester : Manchester University Press.

FURTH, H. G. (1966). A comparison of reading test norms of deaf and hearing children. *Amer. Ann. Deaf,* 111, 461-462.

HAMMERMEISTER, F. K. (1971). Reading achivement in deaf adults. *Amer. Ann. Deaf,* 116, 25-28.

HAMP, N. W. (1972). Reading attainment and some associated factors in deaf and partially hearing children. *The Teacher of the Deaf,* 70, 203-215.

HINE, W. D. (1970). The attainments of children with partial hearing. *The Teacher of the Deaf,* 68, 129-135.

JENSEMA, C. (1975). *The Relationship Between Academic Achievement and the Demographic Characteristics of Hearing-impaired Children and Youth. Report No.* 2, *Series R.* Washington, DC : Gallaudet College, Office of Demographic Studies.

KYLE, J. G. (1977). Audiometric analysis as a predictor of speech intelligibility. *Br. J. Audiol.* (in press).

LEWIS, M. M. (1968). *The Education of Deaf Children: the Possible Place of Finger-spelling and Signing.* London: HMSO.

MOORES, D. F. (1971). *An Investigation of the Psycholinguistic Functioning of Deaf Adults. Research Report No.* 18. Bureau of Education for the Handicapped, Department of Health, Education, and Welfare, US Office of Education.

MOORES, D. F. (1972). Neo-oralism and the education of the deaf in the Soviet Union. *Except. Children,* 38, 377-384.

MOYLE, D., *et al.* (1973). *Readability of Newspapers.* Edge Hill College of Education. (Cited by Bullock, 1975.)

MURPHY, K. P. (1957). *A Survey of the Intelligence and Abilities of 12-year-old Deaf Children.* Ph.D. dissertation. University of Manchester.
MYKLEBUST, H. (1964). *The Psychology of Deafness.* New York: Grune and Stratton.
NORDÉN, K. (1975). *Psychological Studies of Deaf Adolescents.* CWK Gleerup Lund.
OWRID, H. L. (1970). Hearing impairment and verbal attainments in primary school children. *Educ. Res.*, 12, 209-214.
PAIVIO, A. (1971). *Imagery and Verbal Processes.* New York: Holt, Rinehart and Winston.
PROCEEDINGS of the International Congress on Education of the Deaf. Stockholm, 1970.
QUIGLEY, S. P. (1968). *The Influence of Finger Spelling on the Development of Language, Communication, and Educational Achievement in Deaf Children.* Department of Special Education, University of Illinois, Champaign.
RAVEN, J. C. (1960). *Guide to the Standard Progressive Matrices.* London: Lewis.
REDGATE, G. W. (1972). *The Teaching of Reading to Deaf Children: An Experimental Study.* University of Manchester.
VERNON, M. (1976). Communication and the education of deaf and hard of hearing children. In *Methods of Communication Currently Used in the Education of Deaf Children.* London: RNID.
VESTBERG RASMUSSEN (1973). Evaluation of reading achievements of deaf children. In KAMPP, E. (Ed.), *Evaluation of Hearing Handicapped Children.* Denmark: Ebeltoft.
WATSON, T. J. (1967). *The Education of Hearing-handicapped Children.* London: University of London Press.
WRIGHTSTONE, J. W., ARONOW, M. S., and MOSKOWITZ, S. (1963). Developing reading test norms for deaf children. *Amer. Ann. Deaf*, 108, 311-316.

(*Manuscript received* 16*th August*, 1976)

Br. J. educ. Psychol., 47, 149-154, 1977

ANALYSIS OF CHILDREN'S COMPUTATIONAL ERRORS : A QUALITATIVE APPROACH

By J. M. ENGELHARDT
(*Arizona State University, USA*)

Summary. This study was designed to replicate and extend Roberts' efforts at classifying computational errors. 198 third- and sixth-grade pupils were administered an 84-item arithmetic computation test. Those items identified as having incorrect responses were analysed to infer probable student approaches or misconceptions leading to those responses. According to their commonalities, these inferences were clustered to form error types. This procedure resulted in the identification of eight types of errors, two of which corresponded to error types described by Roberts. These were basic fact, grouping, inappropriate inversion, incorrect operation, defective algorithm, incomplete algorithm, identity and zero errors. Finally, examining the distribution of errors according to these eight types led to several tentative generalisations.

INTRODUCTION

Although numerous research studies have been conducted involving arithmetic achievement, little attention has been focused on the qualitative aspects of children's errors, i.e., the types of errors they exhibit. In one attempt to examine pupils' computational errors qualitatively, Roberts (1968) tried to classify errors according to the pupils' methods of attack (called *failure strategies*) which led to the incorrect responses. Using this approach he identified four classes of errors:

Wrong operation: pupil attempts to solve a problem with an inappropriate operation;
Obvious computational error: pupil attempts to solve a problem using an erroneous basic number fact;
Defective algorithm: pupil attempts to solve a problem employing other than basic number fact errors or inappropriate operation errors;
Random responses: pupil attempts to solve a problem in a way showing no discernible relationship to the given problem.

Although Roberts' notion of *failure strategy* is an intriguing construct for describing pupils' computational errors, the class called *random responses* appears to be a catch-all for errors which do not fit into the other error classes. In Roberts' words: " In many instances a response might have been classified as random even though the student may have used a perfectly consistent, if incorrect, strategy " (p. 446). It seems reasonable that these consistent though incorrect strategies might form one or more other error types.

Second, according to Roberts, the *defective algorithm* error type has a number of major subdivisions (e.g., inappropriate reversals, grouping errors, and mixed operations). These subdivisions seem to be sufficiently well-defined to merit consideration as separate error types. A third, and probably the most critical, observation is that no appropriate remedial procedures are inherently suggested by the *defective algorithm* or *random responses* types of errors. One can imagine some appropriate corrective steps if a child performs the wrong operation or employs an erroneous basic fact; but defective-algorithm and random-response types of errors fail to suggest such steps.

Roberts' approach to classifying computational errors appears to have merit as a qualitative way to describe children's computational performance. In light of the questions raised, however, there appears to be a need to replicate and extend Roberts' efforts at classifying computational errors. The present study was, therefore, conducted to meet this need.

METHOD

Sample.

The sample for this study consisted of 198 pupils randomly selected from the third and sixth grades of two public school districts in the greater Phoenix, Arizona, vicinity—71 from a rural district and 127 from an urban district. Pupils in both districts were drawn from five elementary schools and were equally divided by grade and sex. Pupils in the rural district equally represented four ethnic groups—Anglo, Mexican American, Indian and Black.

Procedure.

Within a two-week period, each pupil was individually given a test consisting of 84 arithmetic computation items selected from the *Stanford Diagnostic Arithmetic Test* (Beatty *et al.*, 1966). Administration procedures for this test included noting any unusual behaviours exhibited by the pupils (e.g., finger counting or saying 'plus' in subtraction problems), imposing no time limit and encouraging pupils to complete only as many items for each operation as they felt able to complete. These last two procedures were employed to elicit computational performance which was uncontaminated as much as possible by guessing or the pressures of time.

After the tests had all been administered, the incorrect items were identified and then analysed to determine classes or types of errors. Finally, the distribution of errors by error types among the sample was examined for possible generalisations.

Prior to analysing for types of errors, however, three parameters were accepted. First, Roberts' definition of failure strategies was expanded to include a conceptual component; that is, a group of errors potentially caused by the same misunderstanding would form a separate error type. This expansion was made to avoid the tendency to focus exclusively on the more observable mechanical procedures of computation. Second, since it is possible that a given computational error might be inferred to result from both an erroneous mechanical procedure and a defective concept or principle, it was accepted that such errors might be classified as more than one error type. Although this restriction may have resulted in non-mutually exclusive error types, it was decided that realistic error descriptors were more important than statistical neatness. A third parameter was that broad, high inference, descriptors for error types like ' random ' or ' careless ' were to be avoided.

RESULTS

Error Types.

Within the stated parameters, analysis of pupils' errors led to the identification of eight error types:

(1) *Basic Fact Error*: The pupil responds with a computation involving an error in recalling basic number facts.

An error was classified as a basic fact error if a pupil employed a simple number sentence that is untrue (e.g., $4+3=8$, $63\div7=8$). Such errors were observed most often within the context of multi-digit computations; for example,

$$\begin{array}{r}
\overset{6}{2}7 \\
\times 9 \\ \hline
236
\end{array}$$

(Error: 9×7=56).

(2) *Defective Algorithm*: The pupil responds by executing a systematic (but erroneous) procedure.

Errors of this type could not be described as random responding; indeed, the steps were explainable, and responses to similar computational tasks were predictable. Examples were:

$$\begin{array}{r}
123 \\
\times 42 \\ \hline
186
\end{array}
\qquad\qquad
\begin{array}{rr}
 & 332 \\ \cline{2-2}
32\,) & 996 \\
 & -9 \\ \cline{2-2}
 & 96 \\
 & -9 \\ \cline{2-2}
 & 6 \\
 & -6 \\ \cline{2-2}
 & 0
\end{array}$$

(3) *Grouping Error*: The pupil's computation is characterised by a lack of attention to the positional natural of our number system.

Errors of this type were especially obvious in computational tasks which required ‘ regrouping.’ For example, in the addition problem

$$\begin{array}{r}
57 \\
+93, \\ \hline
1410
\end{array}$$

the columns were often added separately without carrying. Other computational situations not requiring ‘regrouping’ also suggested grouping errors. Problems like

$$3\,\overline{)610}$$

received answers like ‘ 23 r 1, in which the place value of the ‘ 2 ’ was ignored. Sometimes errors which at first appeared to be the result of a defective algorithm could be inferred to be a grouping error. For example, a common multi-digit multiplication error was

$$\begin{array}{r}
13 \\
\times 14 \\ \hline
52 \\
13 \\ \hline
65
\end{array}$$

(4) *Inappropriate Inversion*: The pupil responds with a computation involving the reversal of some critical aspects of the solution procedure.

Computations classified as inappropriate inversions displayed reversals of steps in algorithms which often appeared to promote faster responses. A common example of this was

$$\begin{array}{r}
43 \\
-19 \\ \hline
36
\end{array}$$

where the pupil reversed the subtrahend and minuend in the ones place, thus omitting the borrowing step and deriving a quick response. Another example of this type of error is computations involving carrying in which the place values of the partial sums are reversed. For example

$$\begin{array}{r} \overset{1}{2}3 \\ \times\ 7 \\ \hline 152 \end{array}$$

(5) *Incorrect Operation*: The pupil performs an operation other than the appropriate one.

Examples of this type of error were:

$$\begin{array}{r} 2 \\ \times 3 \\ \hline 5 \end{array} \qquad \begin{array}{r} 13 \\ -\ 1 \\ \hline 14 \end{array} \qquad 4 \div 2 = 8$$

(6) *Incomplete Algorithm* : The pupil initiates the appropriate computational procedure, but aborts it or omits critical steps.

(7) *Identity Errors* : The pupil computes problems containing 0s and 1s in ways suggesting confusion of operation identities.

(8) *Zero Errors*: The pupil computes problems containing 0s in ways suggesting difficulty with the concept of zero.

Pupils who exhibited these last two types of errors appeared either to have inadequate concepts of zero (or one) or to have confused the role of zero and one in the various operations. In the following examples note how the responses would be correct if the zeros were replaced by ones (or *vice versa*) or if the operation was changed:

$$\begin{array}{r} 5 \\ \times 1 \\ \hline 1 \end{array} \qquad \begin{array}{r} 1 \\ -0 \\ \hline 0 \end{array} \qquad \begin{array}{r} 3 \\ \times 0 \\ \hline 3 \end{array} \qquad \begin{array}{r} 0 \\ -5 \\ \hline 5 \end{array}$$

It should be noted that the basic fact and incorrect operation error types directly correspond, respectively, to Roberts' 'obvious computational error' and 'wrong operation' failure strategies.

Distribution of Errors.

The pupils in the sample committed 2,279 errors. On the average, pupils attempted 68·7 of the 84 possible computation items and responded correctly to 57·2, thus committing 11·5 errors. Dividing the sample into quartiles by competence (i.e., total items correct) reveals that although the number of items attempted increases with competence, the number of errors decreases (Table 1). Even though pupils in lower quartiles attempted fewer items, they committed more errors.

TABLE 1

COMPUTATION PERFORMANCE BY QUARTILE.

	(Low)	Quartile		(High)	Total
	1 (N=51)	2 (N=49)	3 (N=49)	4 (N=49)	
	No. Mean	No. Mean	No. Mean	No. Mean	No. Mean
Items Attempted	2415 (47·4)	3193 (65·2)	3902 (79·6)	4097 (83·6)	13607 (68·7)
Items Correct ..	1470 (28·8)	2525 (51·5)	3401 (69·4)	3932 (80·2)	11328 (57·2)
Errors	945 (18·6)	668 (13·7)	501 (10·2)	165 (3·4)	2279 (11·5)

In fact, pupils in the lowest quartile committed over 40 per cent of the total errors. They apparently either attempted problems they were unsure of (regardless of the instructions) or they felt relatively confident about a large number of erroneous concepts and procedures.

The number of errors for each error type is presented in Table 2.

TABLE 2

Distribution of Errors by Type.

	(Low)	Quartile		(High)	Total
Error Type	1	2	3	4	No. (%)
Basic Fact	304	245	218	103	870 (38)
Grouping	159	157	97	90	503 (22)
Inappropriate Inversion ..	211	174	24	63	472 (21)
Incorrect Operation	48	20	19	10	97 (4)
Defective Algorithm	155	164	99	4	422 (18)
Incomplete Algorithm	45	55	48	14	162 (7)
Identity	7	4	6	0	17 (1)
Zero	66	29	43	6	144 (6)

As is evident from the totals, basic fact, grouping, inappropriate inversion and defective algorithm errors were the most common types of errors; in fact, together they account for practically all the errors committed. About one-fifth (408) of the total number of errors were the result of classifying errors as two or more error types, and no errors were identified which could not be classified according to at least one error type.

DISCUSSION

Examining the quartile distribution of errors by type reveals several trends. (*a*) Although (as in Table 1) the number of errors generally decreases across quartiles, basic facts, grouping, inappropriate inversion, defective algorithm, incomplete algorithm and zero errors decline markedly in the upper two quartiles. Apparently the lack of these types of errors is what distinguishes competent performance; conversely, incompetent performance seems to be characterised by a high incidence of such errors. (*b*) In all quartiles basic fact errors were the most common. Increased attention by teachers to mastery of basic facts seems to be an area that would benefit students at all levels of competence. Since the degree of success for instructional methods presently in common use is rather uncertain, there appears to be a need for new instructional approaches; furthermore, research needs to be conducted on those aspects which lead to pupils' difficulty in mastering basic facts. (*c*) The error type which most dramatically distinguishes highly competent performance is the defective algorithm error type. While pupils in the first three quartiles commonly commit defective algorithm errors, pupils in the fourth quartile commit almost no such errors. Apparently one of the more difficult and critical aspects of computation is executing the correct procedure. Although one could speculate that such errors most often occur because these procedures are not meaningful to pupils, research is needed on those aspects which lead to pupils' erroneous computation procedures.

Further examination of the distribution of errors revealed that neither grade level, sex, ethnicity nor urban/rural school district membership suggests the types of errors which pupils commit.

It should be recognised that the inference of error types (i.e., conceptual/procedural approaches to incorrect responses) from pupils' written performance is a limiting factor in this study. Examining only written performance, without the opportunity to investigate a given error further, greatly increased the possibility of misjudging a pupils' erroneous approach. Additional studies need to be conducted in which the inference of pupils' approaches to incorrect responses is based upon clinical interviews. A second limiting factor in this study is the intentional denial of ' careless ' errors. Although the experimental conditions may have artificially masked such errors, no doubt many of the incorrect responses observed in this study were, in fact, the result of temporarily not attending to a given computational task. Future studies, therefore, need to be conducted which deal with ' careless ' errors. The previously mentioned clinical interview technique may be especially helpful in this area. A third limitation of this study was its limited focus upon pupils' approaches to computations which yielded only incorrect responses. It seems reasonable to suspect that pupils sometimes employ erroneous approaches which yield correct responses. Further studies of pupils' erroneous approaches in computation need to address themselves to this area.

Finally, results of this study suggest that the qualitative examination of pupils' computations for types of errors is a viable way to describe their arithmetic performance. Such results further strengthen the argument that teachers should consider more than the number of incorrect responses. Through the examination of pupils' erroneous approaches, teachers are better able to identify and subsequently help them overcome their computational difficulties.

ACKNOWLEDGMENTS—This study was conducted with the support of the University Grants Committee, Arizona State University.

REFERENCES

BEATTY, L. S., MADDEN, R., and GARDNER, E. F. (1966). *Stanford Diagnostic Arithmetic Test (Level II)*. New York: Harcourt, Brace and World.

ROBERTS, G. H. (1968). The failure strategies of third grade arithmetic pupils. *Arithmetic Teacher*, 442-446.

(*Manuscript received* 15*th September*, 1976)

Br. J. educ. Psychol., 47, 155-161, 1977

THE EFFECTS OF PLAY ON MATHEMATICAL CONCEPT FORMATION

By J. ZAMMARELLI and N. BOLTON

(*Department of Psychology, University of Durham*)

SUMMARY. The aim of the study was to assess the effects of a previous play experience on a test of concept formation involving base four arithmetic. Twenty-four children from 10 to 12 years of age were divided into three groups of eight subjects each, matched on teachers' ratings of mathematical ability. A concept formation task was given as an initial measure of ability. Members of group one were permitted total autonomy as to which buttons they pressed on a specially designed toy whose operation embodied base four rules; members of group two were yoked to members of group one so that they observed the same sequence of information; members of group three had no experience with the toy. Each subject was then administered a further concept formation task which involved the same rules existing in the ' toy.' The results showed that group one gained a greater insight into the task and remembered the information more effectively in the short-term. Suggestions for further research are made.

INTRODUCTION

THERE has been much theoretical speculation about the cognitive implications of play but direct, experimental support for the thesis is severely limited. The idea that play experience can lead to the solution of problems is implicit in the Gestalt tradition, for example, in Köhler's (1925) work on insight in apes or Wertheimer's (1961) studies of novel solutions in productive thinking. The importance of the subject's own activity in intellectual development is also stressed in the theories of Piaget (1950), Bruner (1966) and Dienes (1963). However, there are only a few studies which have attempted to substantiate the importance of play for cognitive advance. Among these is the study by Sutton-Smith (1968), which showed that children could think of more unique uses for toys they had played with than for toys belonging to the opposite sex, and a study by Sylva (1976) which demonstrated the effectiveness of play in a Köhler-type insight task. The present study is designed to explore the hypothesis that play can be a useful preliminary experience for the more advanced problem-solving involved in mathematical concept formation.

Since the concept of play appears to overlap with the concept of ' creativity,' (both involving the production of alternative responses without evaluation in relation to a goal), a related area of research is that which has attempted to link tests of divergent thinking with mathematical ability. Bennett (1973), reviewing this literature, points to the contradictory evidence available—Hasan and Butcher (1966), for example, finding a significant relationship between divergent thinking and mathematical ability, whereas Richards and Bolton (1971) found that divergent thinking ability contributed only minimally to performance on a wide range of mathematical tests. However, these latter authors do suggest that further studies should concentrate on exploring the implications of activity methods of teaching mathematics for productive thinking problems of the type described by Wertheimer (1961). The present study pursues this suggestion through an experimental assessment of the effects of active play on the structural thinking involved in mathematical concept formation.

The learning of a mathematical group was chosen as a basis both for the play experience and the concept formation task. This satisfies the necessary conditions as put forward by Dienes and Jeeves (1965) of creating an experimental condition where the subject's first encounter with the data is likely to appear somewhat ' unordered ' and ' chaotic.' It was held that in order to approximate a real life situation, the concept task should be such that the choice of cognitive strategies available to the subject should range from rote memorisation to a single model which has 100 per cent predictive powers. These various strategies must also lend themselves to a scoring process on a scale which can contain several distinct categories. In addition, for our purpose, the mathematical abstractions had to be structured such that they could be represented on a piece of machinery to be used as our pre-task toy. In order to do this, the information contained in the concept task must permit a sequential presentation that can be both controlled by a child, and further, recorded and replicated by a different subject at a later time. By the very nature of mathematical groups, this sort of presentation can be readily accomplished on a toy by relating the push buttons and lights to the rules governing the members of the group, in this case the four colours. The toy was designed to be both portable and easily manipulated, with the children being able to push the buttons at will. It was then a simple matter to obtain three groups of children, a group which played with the toy, a further group who observed the workings of an identical toy that was ' yoked ' to the play-toy, and a third group who had no intervening experience between pre-and post-testing.

Thus, using the methodology developed by Dienes and Jeeves, the aim of this study is to test the prediction that a group which has the benefit of autonomous activity with the specially designed toy will gain greater insight into a related concept formation task than either a group which simply observes the actions of the play group, or one which has no intervening experience with the toy.

METHOD

Sample.

Twenty-four children, aged from 10 years 2 months to 12 years 7 months, were selected from the second and third years of a middle school. For the purpose of forming three groups, matched for mathematical ability, each child was ranked by his mathematics teacher on a 1-12 scale in comparison with his classmates. Table 1 shows that there were no significant differences between the play, yoked and control groups in this respect.

Concept Task.

The concept task which was used is taken directly from the work of Dienes and Jeeves (1965), although coloured blocks were used and not cards. It is known henceforth as the Modulo-4 task and it assesses the child's understanding of base four arithmetic. Wooden blocks of four different colours (yellow, green, red and blue), were used. The procedure can be summarised as follows : The experimenter placed one of his coloured blocks on an area of the table labelled as the centre. The pupil then selected one of his four blocks to place down in his area. As a result of these two blocks, and according to the pre-existing pattern of the numerical group theory, the experimenter then placed a third block on the area of the table designated as resultant, while verbalising the combination, i.e., ' red plus red makes yellow.' The pupil was permitted to view the combination for a few seconds before both the central block and his block were removed, and the resultant block moved over toward the centre to become the new centre colour. The pupil then chose another, or the same, coloured block to place in his area. The sequence was continued until the pupil felt as though he could accurately predict the resultant block for any of the 16 colour combinations. When the pupil desired to check this, a standard verbal test was administered. During the test the children were permitted to manipulate their own blocks but were not permitted to look at the instructor's set which were placed out of sight. If the pupil succeeded in the test, he was then interviewed in order to confirm or contradict the available information as to the method he used to accomplish the task. If the pupil failed, he was reassured that there was no penalty, and the sequence with the blocks continued until he felt ready for another test. There was no limit placed on the number of tests given, and all pupils eventually completed the task.

Scoring of Concept Task.

The number of blocks were recorded that were required by the pupil before he could achieve an 80 per cent success at predicting the 16 combinations of blocks. In addition, two scores were derived from this concept task, both before and after the intervening condition. The first score represents the subject's ability to remember the results of combining any two blocks. It was assessed first by a questionnaire prior to the intervening condition and then by asking a child to complete a four-by-four matrix after the intervening condition : thus in both cases the score is out of a total of 16. The second score was a qualitative evaluation of the child's level of understanding of the concept task. Following completion of the task, each pupil was interviewed and asked how he knew which colour combinations produced the resultant colours. The child's responses were tape-recorded and scored by an independent, experienced rater who had no knowledge of the experimental condition to which the child was assigned. It is these scores which are reported in the text. There was substantial agreement between the two rankings: 32/48 rankings were identical and the other 16 differed by an average of 0·7. Each child was scored on a 1-4 scale both before and after the intervening experience based upon the following criteria. Total memorisation of the block task without any elaboration or awareness of a pattern of relationships of any kind was given a score of ' one.' ' Two ' was awarded when the pupil used a method which relied very heavily upon memory, but which included some basic relationships as aids in acquiring mastery of the system. Such aids which commonly occurred were noticing that yellow plus any colour always yielded yellow (see Figure 1). ' Three ' was scored when the pupil was able to describe the workings of the blocks in terms of a total set of heuristics, that is, several rules and strategies, which when used together could form a systematic approach which would predict any combination of the blocks. In order to attain this category, a score of at least 'two' had to be reached and, in addition, evidence must be available that the blocks were being manipulated in a systematic way that the pupil could express in each and every example presented to him. Usually, this was accomplished by placing the blocks in a special order or arrangement and pointing to them as the procedure was explained. Finally, a top score of four was given only if the pupil was able to explain the workings of the blocks in terms of one rule which could be totally mechanised, that is, demonstrated to work in all cases presented. The complexity of the statement of the rule was not considered crucial, as long as the workings of the blocks were seen as being universally governed by it. As was mentioned before, half scores (such as 1·5) were permitted for a small number of pupils who were clearly between stages or in transition. It was found in a few cases that pupils were able to elevate their score by one-half step during the questioning period, presumably based upon further thought and information acquired there. In these cases the score was given which represented the level attained during the actual experimental task.

FIGURE 1

THE MATHEMATICAL GROUP

	Y	G	R	B
Y	Y	G	R	B
G	G	R	B	Y
R	R	B	Y	G
B	B	Y	G	R

modulo-4 group

Key: Y—Yellow
G—Green
R—Red
B—Blue

The Toy.

The toy had two rows of four buttons each, one horizontal and one vertical. In each row the coloured buttons were arranged in the order yellow, green, red and blue, and the face of the buttons lit when they were pushed to indicate that they were ' on.' A button is turned off by pressing another button in the same row, and each of the two rows, in this respect, were independent of one another. On top of the machine there were four lights, also in the order yellow, green, red and blue. Pressing one button on each of the horizontal and vertical rows caused one of the lights to light according to the rules of the mathematical modulo-4 group.

Procedure.

Each pupil was first given the modulo-4 task. Three groups of pupils were then formed, taking into account year, sex and teacher's ranking of mathematical ability. Pupils were called back to a second session four days later and each was given a sheet of paper which listed the 16 colour combinations and was instructed to remember as many resultants as possible, guessing when unsure. The four original colours of the concept task were present as a set of four blocks during this phase. After completion a child experienced one of the three possible intervening experiences. Pupils in the play condition were given 7½ minutes to play with and explore the toy, pupils in the yoked condition observed the combinations of lights and their resultants, whilst pupils in the control condition had no further experience but returned for testing after 7½ minutes. Finally, each pupil was presented with a chart which contained a list of the four colours vertically and horizontally, thus forming a four-by-four matrix. They were given 16 blocks, four of each colour, and told to complete the matrix. They were asked to explain how they had accomplished this and their explanation was tape-recorded for independent scoring.

RESULTS

The results are presented in Table 1, while Table 2 records the results of applying Mann-Whitney U tests to these data. It may be seen that there are no significant differences between the groups on the four initial measures. On the other hand, on the final concept scores, the differences between the scores of the play group and the two other groups all attain significance, whereas there are no significant differences between the control and yoked groups.

TABLE 1

INDIVIDUAL AND MEAN SCORES ON THE CONCEPT TASK: INITIAL AND FINAL.

Year (2 or 3)	Sex	Teacher's rating of mathematical ability	Initial Concept Scores			Final Concept Scores	
			(*a*) Blocks required prior to solution	(*b*) Memory	(*c*) Conceptual level	(b) Memory	(*c*) Conceptual level
Group 1 Play							
2	M	10	295	4	1	16	3·5
2	M	9	148	8	2	16	3·5
2	F	1	300	12	1	16	4
2	F	7	257	10	1	16	3·5
3	M	11	342	9	1	16	3·5
3	M	4	196	14	2	16	3·5
3	F	12	236	8	1	14	3·5
3	F	9	158	11	2	16	2·5
Mean Values ..		7·87	242	9·50	1·37	15·75	3·43
Group 2 Yoked							
2	M	6	280	12	1	16	2
2	M	5	232	7	1	12	2
2	F	4	180	12	2	12	2
2	F	3	213	15	1	9	1
3	M	7	196	8	1	12	2·5
3	M	10	144	12	2	14	2
3	F	6	194	7	2	11	1·5
3	F	2	210	12	2	12	2
Mean Values ..		5·37	206	10·62	1·5	12·25	1·87
Group 3 Control							
2	M	8	342	12	1	12	2
2	M	2	180	11	2	11	2
2	F	11	368	3	1	6	1
2	F	12	318	8	1	9	1
3	M	8	232	8	1	6	1
3	M	1	274	12	2	16	2·5
3	F	3	126	15	2	16	2·5
3	F	5	254	12	2	9	2·5
Mean Values ..		6·25	262	10·12	1·5	10·62	1·81

TABLE 2

'U' VALUES FROM MANN-WHITNEY U TESTS.

Condition	Comparison			
	Play *v* yoked	Play *v* control	Play *v* non-play	Control *v* yoked
Blocks seen	20·5 NS	27·5 NS	57 NS	18 NS
Memory (pre)........	26 NS	27 NS	75 NS	30·5 NS
Conceptual level (pre)	28 NS	28 NS	56 NS	32 NS
Teacher's rating......	17·5 NS	24·5 NS	42 NS	28·5 NS
Memory (post)	5 P<·002	9 P<·02	14 P<·002	21·5 NS
Conceptual level (post)	0·5 P<·001	1·5 P<·001	2 P<·001	31 NS

DISCUSSION

These results provide a striking confirmation of the hypothesis that play with a specially designed toy can lead to a greater understanding of the rules embodied in a mathematical concept and a better memory for such rules than can be provided by observation of the same stimuli but without manipulation. Indeed, it is only the pupils in the play condition who advanced to higher levels of abstraction, as assessed by the scale of conceptual level. As the same information is available to both 'play' and 'yoked' groups, it is presumably the added autonomy enjoyed by the 'play' group which enabled them to use the information more effectively. A child given play experience can 'programme' his own activities and this may help him to formulate and change his own hypotheses more effectively than the passive recipient of information. In this experiment, therefore, autonomous activity led more rapidly to greater insight and a higher level of understanding.

The instructional opportunities provided in this experiment met Bruner's (1957) demand that instruction should enable the student to "go beyond the information given" and the findings are in accord with the suggestion made by Dienes and Jeeves (1965) that an important feature of this process may be permitting the pupil the opportunity to correct his own false hypotheses. However, although one of the possible educational implications of this experiment is that pupils should be given the opportunity for manipulation of data in play situations, it is probable that the long-term learning process occurring within an educational setting requires the integration of activity methods into more formal teaching. It remains the task of further research using both experimental and observational techniques to suggest how this integration might take place. This research could include measures of long-term memory, since the measure used here might be considered short-term, and assessments of the extent to which play assists in the ability to generalise to other problems and escape the rigid adherence to a set noted by Luchins (1942). It would also be of interest to analyse the strategies of subjects while playing, in order to ascertain which features of this process are most closely related to successful concept learning.

ACKNOWLEDGMENT—This investigation was supported by a grant from the SSRC.

REFERENCES

BENNETT, S. N. (1973). Divergent thinking abilities—a validation study. *Br. J. educ. Psychol.*, 43, 1-7.

BRUNER, J. S. (1957). Going beyond the information given. In BRUNER, J. S. (Ed.), *Contemporary Approaches to Cognition.* Cambridge, Mass.: Harvard Univ. Press.

BRUNER, J. S. (1966). *Towards a Theory of Instruction.* London: Oxford Univ. Press.

DIENES, Z. P. (1963). *An Experimental Study of Mathematics Learning.* London: Hutchinson.

DIENES, Z. P., and JEEVES, M. A. (1965). *Thinking in Structures.* London: Hutchinson.

HASAN, P., and BUTCHER, H. J. (1966). Creativity and intelligence: a partial replication with Scottish children of Getzels and Jackson's study. *Br. J. Psychol.*, 57, 129-35.

KÖHLER, W. (1925). *The Mentality of Apes.* Harmondsworth: Pelican.

LUCHINS, A. S. (1942). Mechanization in problem solving: the effect of *Einstellung*. *Psychol. Monogr.*, 54, No. 248.

PIAGET, J. (1950). *The Psychology of Intelligence.* London: Routledge and Kegan Paul.

RICHARDS, P. J., and BOLTON, N. (1971). Divergent thinking, mathematical ability, and type of mathematics teaching in junior school children. *Br. J. educ. Psychol.*, 41, 32-37.

SYLVA, K. (1976). The role of play in problem-solving. In BRUNER, J., JOLLY, A., and SYLVA, K. (Eds.). *Play: its Role in Development and Evolution.* Harmondsworth, Penguin.

SUTTON-SMITH, B. (1968). Novel responses to toys. *Merrill-Palmer Quart.*, 14, 159-160.

WERTHEIMER, M. (1961). *Productive Thinking.* London: Tavistock.

(*Manuscript received 27th August*, 1976)

Br. J. educ. Psychol., 47, 162-174, 1977

A COMPARATIVE LONGITUDINAL ANALYSIS OF SPECIAL EDUCATION GROUPS

By M. GHODSIAN and M. CALNAN
(National Children's Bureau, London)

SUMMARY. The reading and mathematics performance of five groups of primary school age children (four of which were receiving special education in various forms) were compared longitudinally using a nationally representative sample. The various educational provisions were evaluated taking into account the earlier placement and performance of the children. There was little difference in the reading progress of ESN(M) children in ordinary and special schools although the latter made more progress in mathematics. The results also suggested that children receiving some help within normal schools made most progress.

INTRODUCTION

THERE has been mounting controversy in recent years, both in the UK and USA, about whether children labelled as educationally subnormal should remain in ordinary schools or be placed in ' special ' schools. The arguments generally hinge on the detrimental effects of special schools on the emotional, social or intellectual development of the children attending such schools (Ross *et al.*, 1971 ; Meyers, 1973 ; Miller, 1973). However, there has been surprisingly little systematic research to evaluate the efficacy of these or other types of provisions. The results of the American studies seem to be inconclusive (Hammons, 1972) and the British studies (Stein *et al.*, 1960; Green, 1969; Ascher, 1970; Rutter *et al.*, 1970) also emerge with conflicting results (Tizard, 1966; Pilling, 1973).

This paper presents the results of a longitudinal analysis of the school performance of four groups of children not all of whom were necessarily labelled ESN. However, they are children who at some point (or points) in their school career, before the age of 11, were judged to have educational needs different from those of the majority of children. Some of these children were in special schools, some were receiving help within ordinary schools and others had been recommended for special schools but were in ordinary schools.

Throughout this report, unlike almost all previous studies in this field, account is taken of both earlier performance and placement in comparing the different groups. However, it must be pointed out that in this report we could not take into consideration the whole range of personal and social factors (e.g., motivation, environmental stress) which may differentiate between our groups and directly or indirectly affect their school performance (see Discussion).

METHOD

The data were drawn from the National Child Development Study (NCDS), a multidisciplinary longitudinal study, carried out by the National Children's Bureau, of the physical, social and educational development of all the children in England, Scotland and Wales born in one week of March, 1958. At the ages of 7 and 11 each child underwent a medical examination and completed a number of educational tests, their schools completed a questionnaire, and their parents were interviewed. (See Davie *et al.*, 1972, for a detailed account of the study). On the basis of information from teachers and the local authority medical officers the following groups were delineated separately at 7 and 11 for the purpose of this paper:

(1) Children ascertained as ESN(M) and attending special schools for educationally subnormal children (referred to as ' ESN special school ').

(2) Children ascertained as ESN(M) and attending ordinary schools (referred to as ' ESN ordinary school ').

(3) Children not ascertained as ESN(M) but for whom the teachers said that, irrespective of the facilities in the area, they " would benefit *now* from attendance at a special school " (referred to as ' recommended ').

(4) Children who were receiving help for educational backwardness within the normal school and who were not recommended by their teachers for special school placement (referred to as ' receiving help ').

The above groups are mutually exclusive except for groups 3 and 4: that is, some of the children in the ' recommended ' group were also receiving help for educational backwardness.

The children not falling into any of the above groups comprised the comparison group and were:

(5) ' Normal ' children attending ordinary schools (referred to as ' non SE '—non special education).

The above groups were formed on the basis of information at the two ages. Thus, it must be remembered that there is no information about intermediate placements which might have occurred *between* the two ages.

Children who were also ascertained as having a neuro-physiological defect were excluded as they seem to represent a separate group in terms of etiology, prognosis, etc. (Williams and Gruber, 1967; Zigler, 1966). Nevertheless, there may still have been some with such disorders which were undetected or were not severe enough to lead to ascertainment, who remained in our sample groups.

The measures of school performance used at 7 were:

(*a*) Reading attainment: as measured by the Southgate reading test (Southgate, 1962) (maximum possible score=30).

(*b*) Arithmetic attainment; as measured by a problem arithmetic test constructed for use in the NCDS by the National Foundation for Educational Research (NFER) (maximum possible score=10).

The measures at 11 were:

(*c*) Reading attainment: as measured by a test constructed by the NFER to be parallel with the Watts-Vernon test of reading comprehension (maximum possible score=35).

(*d*) Mathematics attainment: as measured by a test constructed by the NFER for the NCDS and combining a mixture of problems and mechanical items (maximum possible score=40).

Method of Analysis.

The simplest way of detecting change between two or more groups across time is by comparing averages for each group. However, due mainly to the problems of interpreting scores which have been separately transformed and standardised at two ages, this method is unsatisfactory (Fogelman and Goldstein, 1976). The method used by Fogelman and Goldstein (and also adopted here) is designed to identify whether, for children with the same given initial test score, the average at a later stage is different for the various groups they belong to. The problem can be formulated as one of the regression of performance at one age on that of an earlier age for the different groupings of children. This method avoids the above problems and has the additional advantage of providing more detailed information about changes which may occur. Specifically, one can see whether differences between groups at the final stage vary over the range of initial scores.

The problems of measurement error in the 7-year tests has been considered by Fogelman and Goldstein (1976), who concluded that for tests used, the measurement error was small and could reasonably be ignored in the analysis.

RESULTS

Results are presented in two sections. Section (*a*) contains the descriptive statistics for the different groups included in the analysis and section (*b*) presents the findings of the analysis of variance designed to assess the relationship between attainment and the provisions.

(*a*) *Basic results.*

Table 1 presents the numbers of children in each group at 7 and 11.

TABLE 1

NUMBERS OF CHILDREN IN EACH GROUP AT 7 AND 11.

Groups at 7	Groups at 11					
	ESN spcl schl.	ESN ord. schl.	Recom.	Rec. help	Non-SE	Total
ESN spcl schl.	15	2	—	—	1	18
ESN ord. schl.	18	10	2	1	10	41
Recommended	44	23	19	31	108	225
Receiving help	4	6	20	88	449	567
Non-SE	31	48	81	520	11,153	11,833
Total	112	89	122	640	11,721	12,684

The actual analyses were carried out on very slightly different numbers as not all the children in Table 1 had test information at both ages. Table 2 presents the means at 7 for reading and (in parentheses) mathematics, for each category containing four or more children. Table 3 presents the corresponding figures at the age of 11.

TABLE 2

MEANS AT 7 FOR READING AND (IN PARENTHESES) MATHEMATICS FOR GROUPS OF 4 OR MORE CHILDREN.

Groups at 7	Groups at 11					
	ESN spcl. schl.	ESN ord. schl.	Recom.	Rec. help	Non-SE	Total
ESN spcl. schl.	5·0 (0·4)	— —	— —	— —	— —	5·0 (0·4)
ESN ord. schl.	5·8 (0·3)	7·4 (1·5)	— —	— —	11·3 (2·4)	7·3 (1·1)
Recommended	6·5 (0·8)	7·9 (1·3)	9·8 (1·7)	8·5 (1·6)	10·4 (2·7)	9·1 (2·0)
Receiving help	14·0 (3·0)	10·3 (3·0)	12·0 (2·0)	12·8 (2·7)	17·4 (3·6)	16·4 (3·4)
Non-SE	9·8 (1·7)	12·0 (2·1)	12·3 (2·8)	16·8 (3·9)	24·7 (5·4)	24·1 (5·3)
Total	7·4 (1·0)	10·2 (1·9)	11·7 (2·4)	15·8 (3·6)	24·3 (5·3)	23·5 (5·2)

TABLE 3

MEANS AT 11 FOR READING AND (IN PARENTHESES) MATHEMATICS FOR GROUPS OF 4 OR MORE CHILDREN.

Groups at 7	Groups at 11					
	ESN spcl. schl.	ESN ord. schl.	Recom.	Rec. help	Non-SE	Total
ESN spcl. schl.	1·7 (0·1)	—	—	—	—	1·7 (0·1)
ESN ord. schl.	2·3 (0·8)	4·6 (2·3)	—	—	9·1 (3·3)	4·8 (2·1)
Recommended	2·7 (0·3)	5·4 (1·5)	5·7 (1·5)	8·4 (3·7)	8·9 (4·8)	7·0 (3·2)
Receiving help	3·3 (1·0)	6·7 (3·2)	6·4 (3·3)	8·9 (6·6)	12·0 (9·3)	11·2 (8·6)
Non-SE	2·3 (0·6)	6·3 (3·1)	7·6 (4·3)	10·4 (8·0)	17·0 (18·7)	15·7 (17·6)
Total	2·3 (0·5)	5·5 (2·5)	6·9 (3·7)	10·2 (7·6)	16·7 (17·8)	16·0 (16·7)

It can be seen from Table 3 (for example by looking at the means for the row and column totals) that there is a clear gradation in mean scores from the 'Non-SE' to the 'ESN special school' group for both the 7 and 11-year groupings. However, taking the 11-year groups for instance, one difficulty is that we do not know how much these differences are due to difference in initial performance and group membership and how much to the educational provisions of the children at 11.

The focus of this paper will be to compare between groups over particular ranges of scores at 7 years. Table 4 presents the reading and Table 5 mathematics distributions for the 7- and 11-year groups. These are only shown in detail for the range of scores falling roughly below the mean, as this is the effective score range over which the special education children are distributed.

TABLE 4

CUMULATIVE PERCENTAGES OF READING SCORE AT 7 FOR 7 AND 11-YEAR GROUPS.

Reading score at 7	Groups at 7 and 11											
	ESN spcl. schl.		ESN ord. schl.		Recom.		Rec. help		Non-SE		Total	
	7	11	7	11	7	11	7	11	7	11	7	11
0	50·0	12·4	5·4	4·6	4·6	1·7	0·4	0·6	0·03	0·03	0·2	0·2
1–2	50·0	15·3	5·4	6·9	10·2	3·4	0·5	1·2	0·08	0·1	0·4	0·4
3–4	56·3	23·9	24·3	14·9	18·0	7·5	2·1	3·4	0·4	0·4	0·9	0·9
5–6	62·6	42·0	51·3	27·5	31·9	19·1	7·9	7·9	1·4	1·4	2·4	2·4
7–8	75·1	62·0	67·5	47·0	46·6	34·0	16·0	15·9	3·1	3·0	4·5	4·7
9–10	75·1	74·4	71·0	63·1	63·7	50·5	25·2	26·2	5·3	5·1	7·3	7·5
11–12	87·6	86·7	73·7	69·7	75·7	58·8	34·7	37·3	8·3	7·9	10·7	10·8
13–14	93·9	94·3	84·5	78·6	84·0	72·0	43·5	45·0	11·3	10·8	14·0	14·1
15–16	93·9	95·3	87·2	84·3	89·1	79·4	52·6	53·9	14·9	14·4	17·9	18·0
17–19	93·9	98·0	87·2	87·7	93·3	88·5	65·5	68·0	21·6	20·9	24·8	24·8
20–22	93·9	98·0	87·2	90·0	97·5	93·5	76·8	78·8	30·3	29·7	33·5	33·5
23–25	100·0	99·0	87·2	96·9	100·0	97·6	89·3	88·8	42·6	42·1	45·6	45·6
26+		100·0	100·0	100·0		100·0	100·0	100·0	100·0	100·0	100·0	100·0
N=	16	105	37	87	216	121	559	638	11,712	11,624	12,540	12,575

TABLE 5

CUMULATIVE PERCENTAGES OF MATHEMATICS AT 7 FOR 7 AND 11-YEAR GROUPS.

Mathematics score at 7	Groups at 7 and 11											
	ESN spcl. schl.		ESN ord. schl.		Recom.		Rec. help		Non-SE		Total	
	7	11	7	11	7	11	7	11	7	11	7	11
0	81·2	51·0	50·0	23·9	26·0	9·2	6·3	6·3	1·1	1·2	2·1	2·1
1	93·8	70·2	63·2	38·6	47·0	31·7	20·2	19·1	4·6	4·6	6·3	6·4
2	93·8	89·5	78·9	65·9	67·9	58·3	36·1	36·2	13·4	13·0	15·6	15·6
3	100·0	93·3	89·5	84·1	79·5	73·3	53·3	52·2	25·2	24·9	27·7	27·7
4		96·2	97·4	94·3	87·4	88·3	72·5	68·2	39·2	38·9	41·8	41·8
5		99·0	97·4	97·7	95·8	93·3	85·3	79·3	53·3	53·3	55·7	55·7
6+		100·0	100·0	100·0	100·0	100·0	100·0	100·0	100·0	100·0	100·0	100·0
N=	16	104	38	88	215	120	559	638	11682	11582	12510	12532

(*b*) *Analysis of variance results.*

The purpose of the analyses which follow is primarily to take into account the differences among the groups of 11-year-olds which stem from their initial performance and group membership.

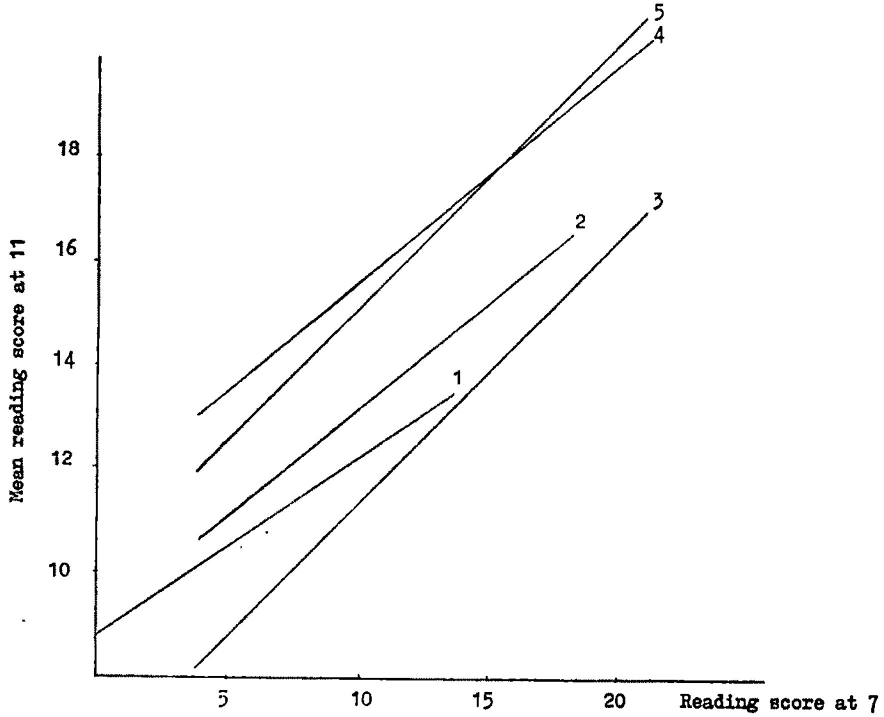
Separate analyses of variance were carried out with reading and mathematics scores at 11 as dependent variables and 7-year score and the groupings at both ages as independent variables. A simple additive ' main effects ' model was used at first, which assumed that the relationship (specifically, the slope of the regression line) between the 11- and 7-year score is the same for each 7- and 11-year group (i.e., there are no ' interactions ' between 7-year score and the groupings at 7 and 11). It also assumes that the differences in average attainment at 11 between, say, the 7-year groups are the same for each 11-year group (i.e., there are no ' interactions ' between the groups at 7 and 11). However, it was apparent in the light of further analyses that this model was inadequate and the results of these analyses are not presented. There were interactions between 7-year score and the groups at each age, and, also, interactions between the groups at the two ages which had to be added to the above simple ' main effects ' model.

Comparison of 11-year groups for each 7-year group:

Appendix A contains the results of those analyses which include the first set of interactions described above, i.e., those between 7-year score and 7- and 11-year groups. Two analyses of variance were carried out; one for reading and one for mathematics. Taking reading, for example, the 11-year score was used as dependent variable with 7-year score, 7-year groupings, 11- year groupings, and interaction terms between 7-year score and the groups at both ages as independent variables. Figures 1 and 2 present, graphically, the results for the group who at 7 were ' non-SE.' As the model includes only the interactions mentioned, the gap between the 11-year groups is the same for each 7-year group and thus the interpretations drawn here apply also to the other 7-year groups, i.e., the special education groups.

FIGURE 1

MEAN READING AT 11 AGAINST READING AT 7 FOR THE ' NON-SE AT 7 ' GROUP. MODEL, INCLUDING ' MAIN EFFECTS ' PLUS 7-YEAR SCORE BY GROUP INTERACTIONS.

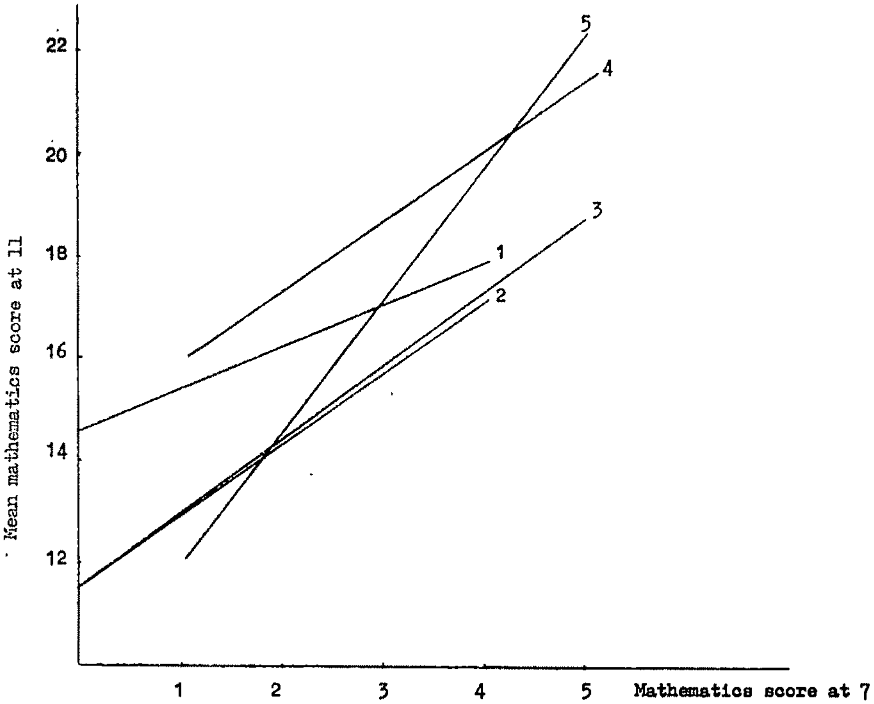


See p. 171 for key to all the figures.

As can be seen from Figures 1 and 2 the regression lines do not cover the same range of 7-year score. This is because, for each group, these lines can only be assumed to be reasonable descriptions of the data over the range where the majority of children score in that group (see Tables 4 and 5). Thus, comparisons can only be made over these ranges.

FIGURE 2

MEAN MATHEMATICS AT 11 AGAINST MATHEMATICS AT 7 FOR THE 'NON-SE AT 7' GROUP. MODEL INCLUDING 'MAIN EFFECTS' PLUS 7-YEAR SCORE BY GROUP INTERACTIONS.



Though these figures do not portray a clear picture the following general interpretations are offered:

Reading: For the same score at 7, the children who were in the ' non-SE at 11 ' or ' receiving help at 11 ' group had similar attainments at 11. Both these groups were doing better than the remaining three (i.e., 'ESN in special school at 11,' ' ESN ordinary school at 11 ' and ' recommended at 11 '). These latter groups perform about equally well (again for the same score at 7) except for the ' recommended at 11 ' group, who do particularly badly at low 7-year scores.

Mathematics: For the same score at 7, the children who were ' receiving help at 11 ' or were in the ' ESN in special school at 11 ' group, do better than the others but are overtaken, in the higher ranges, by the ' non-SE at 11 ' group. Over a considerable range of 7-year score the ' non-SE at 11 ' group perform worse than the ' receiving help at 11 ' group.

The next stage of the analyses was to test the assumption, in the above model, that the differences between the 11-year groups were the same for each 7-year group and *vice versa*. That is, ' interactions ' between the groups at the two ages were added to the above model. For this purpose the two ESN groups at 7 were combined so as to obtain larger groupings. These interactions were not significant for reading (and are not presented) but were so for mathematics. Appendix B contains the results.

The regression lines from this appendix were plotted and the following points can be made:

(i) On the whole the results confirm those of the earlier and simpler model.

(ii) For the same 7-year score, the gap between the ' receiving help at 11 ' groups and the ' non-SE at 11 ' groups (with the latter doing worse) is smaller for children who had been ' non-SE ' at 7 than for other 7-year groups.

(iii) For the ' receiving help at 7 ' or ' recommended at 7 ' groups the ' receiving help at 11 ' do particularly well by being consistently superior (for the same 7-year score) to all the other 11-year groups.

For these same 7-year groups, the ' ESN special school at 11 ' children do better (for the same 7-year score) than those in the ' non-SE at 11,' ' recommended at 11 ' or ' ESN ordinary school at 11 ' groups over a wider range (compared to the other 7-year groups) of 7-year score.

Comparisons of 7-year groups for each 11-year group.

In this section the 7-year groups are compared for each 11-year category. The very restricted range of 7-year score (Tables 4 and 5) obtained by the two 7-year ESN groups, though not surprising, make any interpretations relating to these very tentative, and applicable over narrow ranges. The small size of the groups in this section further restrict the interpretations. These comments which apply particularly in the case of ESN children who were in special schools at 7, must be borne in mind in reading the following.

Reading: The regression lines from Appendix A were plotted. Figure 3 presents the reading results for the ' ESN special school at 11 ' group. The following descriptions apply to the other 11-year groups as well: the ' recommended at 7 ' group do better, at 11, than the children in the ' ESN ordinary school at 7,' ' receiving help at 7,' ' non-SE at 7 ' and ' ESN special school at 7 ' groups for the same 7-year score. These groups do progressively worse in the above order. The gap between the groups tends to decrease with increasing 7-year score.

Mathematics: In this case the results proved to be extremely complex. Not only were the pictures presented by including the first and the second set of interactions complex, they also differed from each other substantially in some cases. In view of this only the results for the two ESN groups at 11 are discussed as they are probably of most importance and interest. Further, the results of the two models described above (the first model consisting of the ' main effects ' plus 7-year score-by-group interactions and the second being the first model plus group-by-group interactions) will be discussed in conjunction with each other.

FIGURE 3

MEAN READING AT 11 AGAINST READING AT 7 FOR THE 'ESN SPECIAL SCHOOL AT 11' GROUP. MODEL, INCLUDING 'MAIN EFFECTS' PLUS 7-YEAR SCORE BY GROUP INTERACTIONS.

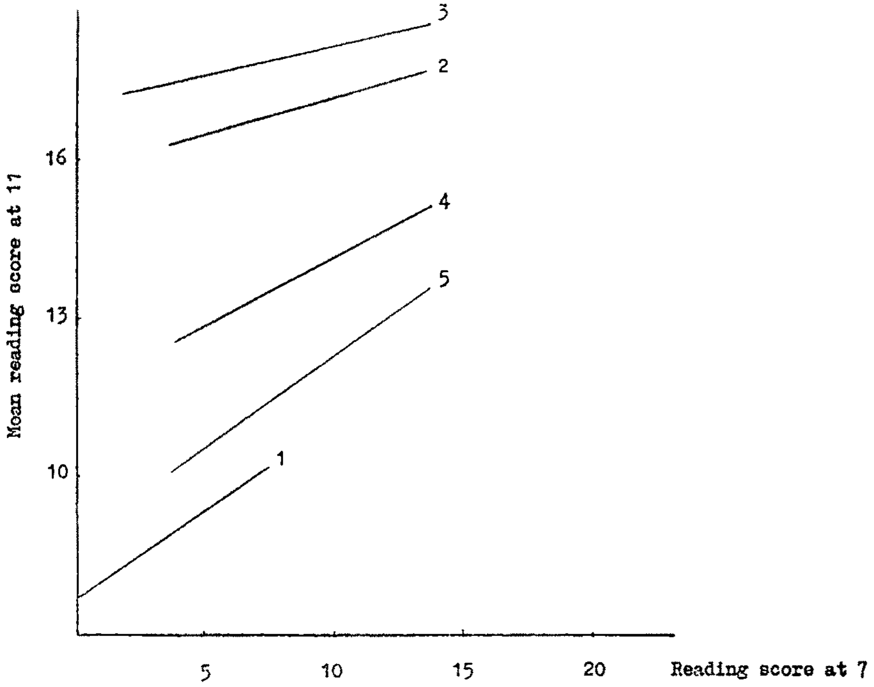


FIGURE 4

MEAN MATHEMATICS AT 11 AGAINST MATHEMATICS AT 7 FOR THE 'ESN SPECIAL SCHOOL AT 11' GROUP. MODEL INCLUDING 'MAIN EFFECTS' PLUS 7-YEAR SCORE BY GROUP INTERACTIONS PLUS GROUP BY GROUP INTERACTIONS.

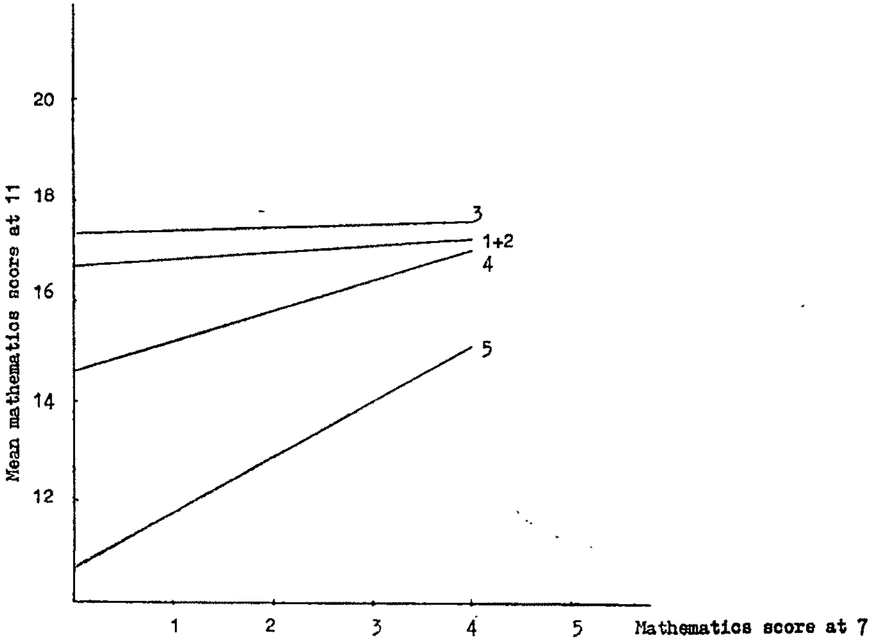


Figure 4 presents the results for the children who were in the ' ESN special school at 11' group. The results are similar for the other ESN group, i.e., those in ordinary schools at 11. The following points can be made: for the same 7-year score, the ' ESN at 7 ' and ' recommended at 7' groups do equally well and both do better than the ' receiving help at 7 ' and (especially) the non-SE at 7 ' groups. The evidence from Appendix A (i.e., the model not including group by group interactions) indicates that the ESN children in ordinary schools at 7 do much better than those who were in special schools. The gap between the 7-year groups generally narrowed with increasing 7-year score.

LEGENDS TO FIGURES

(*a*) All graphs are only shown for the lower part of the 7-year-scores.
(*b*) Nomenclature:
  (1) Signifies the ESN special school group.
  (2) Signifies the ESN ordinary school group.
  (3) Signifies the recommended group.
  (4) Signifies the receiving help group.
  (5) Signifies the non-SE group.

DISCUSSION

In this section an attempt is made to draw together the rather fragmentary results presented so far.

Reading: Among children whose attainment had been comparable and low at 7, the ESN children in special and ordinary schools at 11 do about equally well at the later age (those in ordinary schools doing slightly better at higher scores). However, the type of school which they had been attending at 7 seems to be quite important. Those who were ESN and had been in ordinary schools at 7 made better progress than those who had been in special schools at 7. Perhaps what is important here is that those children who were in some way ' receiving attention ' within the ordinary school at 7 (i.e., by receiving help, being recommended or being ascertained ESN) were, for the same 7-year score, doing better at 11 than the two extremes (i.e., those who, at 7, were ' non-SE ' and those who were ESN in special schools).

Mathematics: The pattern of results almost always seemed to be more complex for mathematics than for reading. This might be partly due to the fact that mathematics may be a more diverse task than reading, thus making it perhaps more difficult to tap a common mathematics ability. Another reason might be the limitations of our data (namely, small group sizes and using tests designed for the general population which lack sensitivity in the lower ranges with which we are concerned). Nevertheless, some overall remarks can be made.

For the same score at 7, ESN children in special schools at 11 do better than ESN children in ordinary schools. There does again seem to be tendency for those who were recommended, ESN in ordinary schools or receiving help (i.e., those receiving some ' attention ') at 7 to do better at 11, for the same 7-year score, than the two extremes (non-SE and ESN special school at 7).

What is particularly impressive is that for the same score at 7, for both reading and mathematics, on the whole those receiving help at 11 do better than other 11-year groups especially for low 7-year score. Though it is not known for how long the children were receiving help this does seem to indicate that receiving special help in the ordinary school without being ascertained or placed in special schools is beneficial. There is additional support for this proposition from the fact that in mathematics for the same score at 7, the children who were receiving help both at 7 and 11 do better than *all* other groups. In the case of reading this applies to those who received help at 7 and were either ' non-SE ' or ' receiving help ' at 11. (All this of course only applies in the ranges considered above.)

Further, the fact that ESN children in special schools at 7 make less progress than others by 11 argues against early placement in special schools. At the same time the poor relative performance of those who were ' non-SE at 7 ' argues for early detection of need.

While research of the type reported here is badly needed to establish, at a general level, the relationship between attainment and different provisions it would be more valuable if accompanied by investigations into the reasons and mechanisms giving rise to the results. Thus, we do not really know why for given initial performance the ESN children in ordinary schools do about as well in reading as those in special schools, but perform less well in mathematics. This is unlikely to be due to different methods of teaching as the children were distributed nationally. Could it be due to a general bias in emphasis towards one subject?

Nor do we know why those who were at a special school or were ' non-SE ' at 7 performed relatively worse than others at 11. Perhaps we can make some obvious reasoned guesses about the latter group, i.e., " they did not receive enough appropriate attention at an earlier age when they ' needed ' it." But why is being at a special school at 7 a disadvantage at 11? Is it a consequence of the ' stigma ' of going to special school? Is it a function of the school organisation? Is it that in early years these schools emphasise other aspects of development?

These questions lead us to two further and crucial points. The first is that the *actual* criteria used (or the factors involved) in allocation to different provisions are not clearly known and may vary from one local authority to another (Presland, 1970; Williams, 1965). It might well be that the differences obtained are due to the fact that particular types of children with particular circumstances are allocated to each provision.

This brings us to the second point which bears on the limited scope of studies which just investigate scholastic achievement. There is a growing awareness of the cultural relativity associated with ' diagnosing need ' in this area (Mercer, 1968). While this concept can easily become an excuse for ' non-action ' it has the merit of drawing attention to the child's total life experience and social background. (See, for example, Zigler, 1966, for a discussion of experiential and motivational factors). In addition, there is the awareness of environmental ' handicaps ' or ' deficiency ' as a concomitant if not the cause of ' subnormality.' While, again, there is a danger of confusing ' deficiency ' with ' relativity ' this warns against an emphasis on schooling when the problem is ' societal ' : witness the remarkable difference in social class distribution between ESN children with and without organic defects (Williams and Gruber, 1967; Stein and Susser, 1960). However, all this is not to detract attention from those relatively few cases (Rutter *et al.*, 1970) who display organic deficiencies.

Considerations such as the above plus the complexity of the results presented here, show the futility of ' all or nothing ' solutions to the debate about special education. Instead, they underline the need for much more subtle distinctions to be made if evaluation, let alone innovation of provisions, is to be attempted (Keogh, 1975).

ACKNOWLEDGMENT—We are grateful to Harvey Goldstein for carrying out the statistical analyses. We thank other members of staff of the National Children's Bureau, in particular Ken Fogelman, for their help. This research was financed by grants from the DHSS, the DES and SSRC which are gratefully acknowledged.

REFERENCES

ASCHER, M. A. (1970). The attainments of children in ESN schools and remedial departments. *Educ. Res.*, 12, 215-219.
DAVIE, R., BUTLER, N., and GOLDSTEIN, H. (1972). *From Birth to Seven.* London: Longman in association with National Children's Bureau.
FOGELMAN, K., and GOLDSTEIN, H. (1976). Social factors associated with changes in educational attainment between 7 and 11 years of age. *Educ. Studies*, (2, 95-109).
GREEN, L. F. (1969). Comparison of school attainments. *Spec. Educ.*, 58, 9-12.
HAMMONS, G. W. (1972). Educating the mildly retarded: a review. *Except. Child.*, 38, 565-570.
KEOGH, B. K. (1975). Social and ethical assumptions about special education. In WEDELL, K. (Ed.), *Orientations in Special Education.* London: John Wiley and Sons.
MERCER, J. (1968). Mental subnormality. In RUBINGTON, E., and WEINBERG, M. S. (Eds.), *Deviance: the Interactionist Perspective.* London: Collier-Macmillan.
MEYERS, C. E. (1973). The school psychologist and mild retardation—report of an *ad hoc* committee. *Mental Retardation*, 2, 15-20.
MILLER, P. (1973). All children are special. *Assoc. educ. Psychol. J.*, 3, 40-46.
PILLING, D. (1973). *The Handicapped Child: Research Review*, Vol. III. London: Longman, in association with National Children's Bureau.
PRESLAND, J. (1970). Who should go to ESN schools? *Spec. Educ.*, 59, 11-16.
ROSS, S. L., DE YOUNG, H. G., and COHEN, J. S. (1971). Confrontation: special education placement and the law. *Except. Child.*, 38, 5-12.
RUTTER, M., TIZARD, J., and WHITMORE, K. (1970). *Education, Health and Behaviour.* London: Longman.
SOUTHGATE, V. (1962). *Southgate Group Reading Tests: Manual of Instructions.* London: University of London Press.
STEIN, Z., SUSSER, M., and LUNZER, E. A. (1960). Reading reckoning and special schooling among the mentally handicapped. *Lancet*, 2, 305-307.
STEIN, Z., and SUSSER, M. (1960). The families of dull children: a classification for predicting careers. *Br. J. prev. soc. Med.*, 14, 83-88.
TIZARD, J. (1966). Schooling for the handicapped. *Spec. Educ.*, 55, 4-7.
WILLIAMS, P. (1965). The ascertainment of educationally subnormal children. *Educ. Res.*, 7, 136-146.
WILLIAMS, P., and GRUBER, E. (1967). *Response to Special Education.* London: Longman.
ZIGLER, E. (1966). Mental retardation; current issues and approaches. In HOFFMAN, M. L., and HOFFMAN, L. W. (Eds.), *Review of Child Development Research.* Vol. 2, New York: Russell Sage Foundation.

(*Manuscript received* 21*st June*, 1976)

APPENDIX A

Analyses of variance findings using model with 11-year score as dependent and 7-year score, groups at 7, groups at 11 and interactions between 7-year score and groups at both ages as independent variables.

Independent Variables	Dependent Variables			
	11-year reading score Mean=16·11 Variance=38·72 N=12619		11-year mathematics score Mean=16·87 Variance=106·08 N=12510	
Overall Constant	4·46		2·08	
Education at age 7	Intercept	Slope	Intercept	Slope
ESN special school	−1·43	·28	−·90	1·15
ESN ordinary school	·51	·07	·45	−·59
Recommended	·99	·04	−·54	−·28
Receiving help	·20	·18	−·41	·48
Non-SE....................	−·24	·28	1·40	1·15
Education at age 11				
ESN special school	−2·71	·07	−1·58	− ·35
ESN ordinary school	− ·75	·14	−1·31	·28
Recommended	−1·03	·24	−1·17	·32
Receiving help	1·69	·15	1·31	·21
Non-SE....................	2·77	·25	2·75	1·45
Residual Mean Square	22·60		67·08	

All variables are significant at the ·001 level.

APPENDIX B

Analysis of variance findings using model with 11-year mathematics score as dependent variable and 7-year score, groups at 7, groups at 11, interactions between 7-year score and groups at both ages and interactions between 7 and 11-year groups as independent variables.

Dependent variable=11-year mathematics score, Mean=16·87, Variance=106·08 (N=12510), Overall constant=2·19.

Education at age 7	ESN special school at 11		ESN ordinary school at 11		Recommended at 11		Receiving help at 11		Non-SE at 11	
	Inter-cept	Slope	Inter-cept	Slope	Inter-cept	Slope	Inter-cept	Slope	Inter-cept	Slop
ESN	−1·58	·14	−·50	·73	·54	·83	3·74	·66	−1·76	1·77
Recommended	−1·49	·07	−·81	·66	− ·95	·76	1·44	·59	− ·45	1·70
Receiving help ..	−1·35	·62	−·90	1·21	− ·32	1·32	2·93	1·14	1·22	2·25
Non-SE........	−2·61	1·14	−1·76	1·73	−1·26	1·83	1·62	1·66	4·23	2·77
Residual Mean Square=66·94										

All variables are significant at the ·001 level.

NOTE—The relationship between 11-year mathematics score and 7-year arithmetic score is adequately described by a straight line, whereas that for reading becomes markedly non-linear for high 7-year scores. This, at least in part, reflects the fact that the 7-year reading test discriminated poorly between good readers and resulted in nearly one-third of children obtaining the top scores of 29 and 30. In order to simplify the analysis, the scale of the 7-year reading score has, therefore, been transformed to give an approximately linear relationship.

Br. J. educ. Psychol., 47, 175-186, 1977

SYMBOLIC PLAY: THE ROLE OF SUBSTITUTIONS IN PRETENCE AND PUZZLE GAMES

BY CLAIRE GOLOMB
(*University of Massachusetts at Boston*)

SUMMARY. In a study of play activities, 60 middle-class children, ages 2·8—5·8 were tested on six object substitution tasks. Of these, three dealt with puzzle games and three with pretence games. On both sets of tasks the substitute objects ranged from suitable to highly unsuitable ones. The results indicated that object substitution on the pretence games was more extensive than the substitution of parts on the puzzle games. The selections on both games, however, were rule-governed and orderly in terms of the sequence and type of object selection, beginning with the most suitable ones and declining when only unsuitable ones were available.

INTRODUCTION

TRADITIONALLY, most students of child play have conducted their investigations under more or less naturalistic conditions, recording play behaviour in the home, the nursery school, the playground, or the child-therapist's office. The major accounts of child play have stressed its psychodynamic significance and emphasised the affective value of symbolic or pretence play for the child's emotional well-being. This holds also true of Jean Piaget's developmental analysis of play which falls squarely within this tradition. It is Piaget's (1962) contention that in symbolic play the child subordinates reality to the whims of the ego, i.e., distorts reality and transforms it to fulfill his needs and desires. According to this view, symbolic play is symptomatic of the child's cognitive egocentricity and conceptual immaturity. As evidence Piaget cites the child's playful tendency arbitrarily to substitute objects in play and to behave as if anything can stand for anything. He maintains that the child's substitutions lack rhyme or reason, and represent essentially idiosyncratic and undisciplined forms of thought.

Recently, the possibility that such activities may serve a cognitive function has been raised. Several authors have stressed the transformational nature of play and its potential usefulness for the acquisition of representational sets and conceptual thought, skills which are essential for later scholastic achievement (El'Konin, 1966; Feitelson and Ross, 1973; Smilansky, 1968; Sutton-Smith, 1967). This interest in the cognitive aspects of symbolic play has led to several play training studies which demonstrate that symbolic play behaviour can be enhanced by various modelling techniques. However, the analysis of the symbolic transformation of the object in play has been neglected. What in fact do we know about the nature of these substitutions? Piaget assumes that object substitutions are a function of spontaneous developmental trends which begin during the latter part of the second year of life. The early substitutions are arbitrary, undisciplined symbolic transformations and are only gradually replaced by orderly and conceptually regulated ones. El'Konin (1966) and Smilansky (1968) disagree with this position and maintain that the symbolic transformational process requires adult modelling and is largely taught by the adult. Thus, the issue of the young child's capacity to pretend that a missing object can be represented by another, though dissimilar one, is of some importance.

The available empirical data, however, are of limited scope. One of the few studies dealing with this problem was conducted by Sarah Sliosberg (1934). Sliosberg differentiated between child play in a predominantly realistic context and play in the context of pretence. She studied pre-school children's willingness to substitute objects during pretence play and during play with art materials and block building. Under these two contrasting conditions she discovered important differences in the child's selections, differences which she attributed to the effects of the tasks and the materials. These findings seemed to suggest that the child's behaviour in the constructive play activities (reality context) was subject to certain implicit rules of the reality oriented game. Her data, however, are not altogether clear, and the lack of statistical analyses prevents us from determining whether an obtained difference was significant. While she found a tendency to substitute objects more readily in the pretence situation than in the realistic one, this was not an altogether consistent trend. Furthermore, she failed to analyse fully the nature of the realistic and of the pretence situation, and ignored the ready transformation of one into the other as a function of examiner expectations. Although she mentioned the existence of ' rules ' and of ' limits ' in the realistic situation, she failed to analyse these factors in the case of pretence.

In view of the conflicting conceptions of the nature of the child's pretence in symbolic play, the present study was designed to elucidate children's playful object substitutions. Pre-school children were presented with a set of three pretence games and a set of three puzzle games. In each case the game required the selection of an object or puzzle part in order to fulfill the requirements of the task. The available objects varied from highly suitable to highly unsuitable ones. This design permits the investigation of (*a*) the child's spontaneous responses when substitute objects are made available; (*b*) the effect of suggestions on the child's selections; and (*c*) the limits of acceptable object substitution.

METHOD

Sample.

Sixty middle-class pre-school children (30 males and 30 females) who were enrolled in a university nursery school or day care centre participated in both study 1 and 2. The children were classified according to age: group 1 consisted of 16 children ranging in age from 2·8 to 3·7 with a median age of 3·0; group 2 consisted of 25 children ranging in age from 3·8 to 4·7 with a median age of 4·0; group 3 consisted of 19 children ranging in age from 4·8 to 5·8 with a median age of 5·0.

STUDY 1

Puzzle Play Tasks and Procedures.

Each child was asked whether he knew what a puzzle was, had played with puzzles before, and noticed that puzzle pieces had to fit well together. Next each child was presented with three puzzle tasks, one at a time and in random order; a Manikin puzzle, a Car puzzle, and an Elephant puzzle. Each puzzle was presented in its near completed form with a *single* part missing. In the case of the Manikin—the head, in the case of the elephant—the back, in the case of the car—the rear wheel. In each presentation the missing part was displayed among an array of pieces which varied in terms of degree of fit to the standard missing part. The child was encouraged to look over the collection of pieces, and asked to find the one that would best complete the puzzle.

The three puzzle tasks were administered individually and in a single session.

(*a*) *The Manikin Puzzle.*

Materials. The Manikin consisted of six parts: head with facial features, torso, arms and legs. The Manikin was presented with the head (the standard part) missing and displayed among the following array: a blank face (identical with the standard part minus the facial features), an enlarged face (three times the standard size), a small face (one-third the size of the standard), a square face, a cat's face, a large oval (approximately three times the standard size), a medium oval (matched in size to the standard part), a small oval (one-third of the size of the standard), a disc, a hand, a telephone, and a bowl of fruit.

Instructions. " Here is a puzzle of a man. You see the head is missing. These are different pieces (examiner pointed to the array) ; see if you can find the right part to finish the puzzle man." After the child made his selection and inserted the part to complete the manikin, the examiner continued: " Fine! Now I take this part away; look well and tell me if there is another part that you could use to finish the puzzle." The examiner urged the child to look over the array. If the child merely pointed to a part, he was encouraged to pick the piece up and place it where it belonged in the puzzle. The examiner also questioned the child's choices, asking him/her: " Is this a good part? Does it make a good puzzle-man ? " The experiment continued until the child stated that he could not find a suitable part to the puzzle.

(*b*) *The Car Puzzle.*

Materials. The car puzzle consisted of three parts : front part, rear part, and a rear wheel cut-out. The puzzle was presented without the rear wheel (the standard part) which was displayed among the following array: a standard part with an incongruous facial drawing, a wheel matched to the standard wheel but not to the whole cut-out, a large wheel, a small wheel, a square wheel, a face, a steering wheel, bicycle handlebars, a cup, a clock, a boot, and a flower.

Instructions. " Here is a puzzle of a car. You see the wheel is missing. These are different pieces (examiner pointed to array); see if you can find the right part that will finish the car puzzle . . . " (same procedure as for Manikin).

(*c*) *The Elephant Puzzle.*

Materials. This puzzle was composed of three parts: head and trunk, belly and legs, and back. The elephant was presented with the back (standard part) missing. This part was displayed among an array including the following parts: standard part with an incongruous drawing, an oversized back (twice the standard), a small back (half the size of the standard), a rectangle, a triangle, and a disc approximately matched in area to that of the standard, a large rectangle (twice the size of the standard), a cat's face, a cut-out of milk-bottles, a tree, and a driver.

Instructions. " Here is a puzzle of an elephant. You see the back is missing . . . " (same procedure as for Manikin).

Data Analysis.

The whole session was tape-recorded and transcribed verbatim. The complete protocols consisted of the transcribed tape-recordings and the written records of the examiner. The data were analysed for type, incidence, and order of substitute selections. Two independent scorers who were unfamiliar with the subjects scored all the records; percentage of agreement was as follows:

Manikin—93 per cent; Car—93 per cent; Elephant—94 per cent.

RESULTS

The overall results demonstrate the orderliness of the child's selections which were determined by a rule of ' good fit,' i.e., the chosen part had to fit exactly into the puzzle.

In general, the first choice was the most appropriate one. Those youngsters who could be induced to continue the substitution process proceeded, in a somewhat playful manner, from the most suitable parts (the standard part) to relatively neutral ones before incongruous choices were attempted. This held true for all three age groups and the three puzzle tasks and was reflected in the order of selection, the type of substitution, and the number of rejections. Those children who were willing to go beyond the optimal choice (standard part) made a clear distinction between a ' good fit ' and ' something that could be used.' All children indicated their awareness of the absurdity of the later selections, and even for the most compliant ones there were limits to the acceptability of substitutes.

Task Effects.

In all three puzzles a sharp decline was noticed in the selection of parts once suitable items had been eliminated, a decline which was affected by the structure of the particular puzzle. The effect of the structural differences could be seen in the variable number of ' complete ' selections made by our young participants, namely, the number of children who selected all available parts of a given puzzle array. In the case of the Manikin puzzle—nine out of 60 children selected all available parts; in the case of the Car puzzle—seven children selected all available parts ; and in the case of the Elephant puzzle—only four children made complete selections. For the distribution of selected parts, see Table 1.

The location and type of the cut-out affected the extent of substitute selections. Thus, for example, the Manikin puzzle allowed far greater flexibility in the selection of parts since the missing standard part, the head, did not have to fit precisely into an enclosure and could merely be joined to the top of the figure. In this puzzle the selections were more humorous than in the other puzzles (see Qualitative Observations section), beginning with the animate pieces before passing to neutral and incongruous ones. The meaning of a part was the over-riding factor which determined its suitability in conjunction with its form and size.

The selections for the Car puzzle also showed a radical drop from the optimal fit (the standard wheel part) to other types of substitute parts. Form was the determining factor in the selection process and guided the search for the correct insert. After the standard cut-outs had been eliminated, the rule of ' roundness ' determined the selections, followed by very few incongruous selections, i.e., a high percentage of rejections (see Table 1).

Form was also an important factor determining the selection of parts for the Elephant puzzle, and again the notion of precise fit dominated the responses. The selections began with the standard back part, continued with the over- and under-sized back parts, followed by the selection of parts which were approximately matched to the area of the standard, and tapered off with only a few incongruous selections.

Altogether, the order of selection closely followed the principle of object congruency.

TABLE 1

DISTRIBUTION OF SELECTION OF PUZZLE PARTS IN PERCENTAGES FOR 60 PRE-SCHOOL CHILDREN

Manikin Puzzle												
Standard Face	*Standard without Features*	*Standard Large*	*Standard Small*	*Cat's Face*	*Square Face*	*Oval Medium*	*Oval Large*	*Oval Small*	*Disc*	*Fruit*	*Telephone*	*Hand*
95	65	55	48	52	45	47	30	28	33	28	25	25
Car Puzzle												
Standard Wheel	*Standard with Inc.* Drawing*	*Disc Matched*	*Disc Small*	*Disc Large*	*Face*	*Square Wheel*	*Steering Wheel*	*Handle-Bars*	*Clock*	*Cup*	*Boot*	*Flower*
100	45	42	35	33	30	27	28	23	23	20	13	13
Elephant Puzzle												
Standard Back	*Standard with Inc.* Drawing*	*Standard Large*	*Standard Small*	*Rectangle Matched*	*Rectangle Large*	*Triangle*	*Disc*	*Cat's Face*	*Driver*	*Tree*	*Milk Bottles*	
95	67	32	30	30	18	25	22	25	18	18	17	

*Inc. = Incongruous

Age Effects.

The order of selection of substitute parts was unaffected by the age of the children. Younger children (ages 2·8—3·7), however, made more incongruous selections than older ones (3·8—5·8). This was a consistent finding for all three puzzles, though it was somewhat less pronounced for the Manikin puzzle. Sex differences were not observed.

Qualitative Observations.

The protocol material clearly revealed the child's tendency to make only few and fitting selections. Faced with the very suggestive instructions given by the examiner, the children gave clear indications that they differentiated between ' optimal fit ' and ' what might be used.' The first optimal choice was always followed by some show of hestitation before verbal compromises were attempted. These compromise solutions took various forms. On the Manikin puzzle, for example, when an incongruous part was substituted for the head the following combinations, usually in hyphenated form, were offered: ' telephone-man,' ' baby-man,' ' little-man,' ' little-fellow,' ' halloween-mask,' ' wolf-man,' ' cat-man,' ' spider-man,' ' Santa,' ' apple-man,' ' fruit-man,' ' astronaut.' Similar solutions were offered for the Car and Elephant puzzles.

Other compromise solutions which clearly reflected a degree of conflict were expressed as follows; " Does not fit very well but there are no more parts," " cat-head is funny but can fit," " this can cover the elephant, sometimes an elephant has a cover " (refers to the standard with the incongruous drawing), " cat-part fits because it can travel on the elephant's back," " more a cat than a man," "it really is a telephone, but it's okay for the man," "someone will eat his head off " (fruit part).

Some children made substitute selections but described the deficiencies of each part in great detail or reaffirmed the original function of the part: " it really is . . . " Others suggested cutting off the edges of the substitute to create a better fit.

Finally, several children opted for a composite solution, combining several pieces to fill the area of the cut-out while simultaneously doing justice to the representational aspect, for example, filling the cut-out of the car with a square piece on top of which a wheel was placed.

STUDY 2

Pretence Play Tasks and Procedures.

Each child faced a revolving stage, partitioned into three sections which permitted three different games to be played. The experimenter defined each game as a pretence game and freely encouraged object substitutions. The following games were played: ' Feeding the Hungry Baby,' ' Going to the Beach,' and the ' Petshop.'

The three pretence games were administered individually, in a single session, and the order of presentation was randomised.

(*a*) *Feeding the Hungry Baby.*

Materials. The stage displayed a doll which the experimenter described as a hungry baby. The stage background, defined as a kitchen, exhibited the following objects, dispersed in random order: three bottles (a doll-sized bottle, a baby-bottle, and a quart-sized bottle), a can of pineapple, a picture of food, a cardboard roll, a blob of Plasticine, a wooden block, two candles (one small, one large), a pencil, a furry toy-animal (rabbit), a toy car, and a hairbrush.

Instructions. " Let's pretend this is a baby (examiner pointed to the doll) that just woke up. It is hungry and wants to eat. This (examiner pointed to the stage) is your kitchen. Can you find something in your kitchen to feed the baby? " After the child made his selection, the examiner removed the object out of sight and continued: " Your baby is still hungry, can you find some more food ? " The game was continued until all objects had been chosen or the child refused to make further selections. After the first refusal the examiner questioned the child: "Are you sure that there is no food in the kitchen? Please, look carefully." If the subject seemed to pick objects at random or without verbally indicating the nature of the object, the examiner questioned the child : " Can you tell me what this (pointed to the chosen object) is? "

(*b*) *Going to the Beach.*

Materials. The stage displayed a picture of children playing on a beach. The following objects were randomly arranged in front of the stage: a small doll, a picture of a child, a smilebutton, a toy giraffe, a photograph of a tadpole-man, a stuffed nondescript form, a Plasticine stick, a cardboard roll, a string, a pencil, a carrot, a cookie, a key purse, a shoeboot.

Instructions. " Let's pretend it is a sunny day and you and your friend are going to the beach. So now take your friend along and let's go to the beach." After the child indicated his choice, the examiner assigned a name to the object, removed it from sight, and continued : " Your friend, Johnny has to go home, please take your friend Jackie along." This procedure continued until all objects had been chosen or the subject refused to make further selections. If the choice seemed automatic or random, the examiner pointed to the chosen object and questioned the child. " What is this? " The examiner alternated male and female names for the friend.

(*c*) *The Petshop.*

Materials. The following objects were displayed in random order: a stuffed toy kitten, a porcelain kitten, a picture of a kitten, a stuffed toy dog, a stuffed miniature porcupine, a stuffed nondescript form, a Plasticine blob, a wooden block, a furry slipper, a carrot, a small doll, a toy cradle.

Instructions. " Let's pretend today is your birthday. Let's pretend you always wanted a kitten for a pet. Now we are on our way to the petshop. Here is the petshop; you may choose a kitten for your pet." After the child made his choice, the examiner removed the object and stated: " This kitten is sick and has to go to the doctor. Can you find another one? " If the child's response was negative, the examiner questioned him: "Are you sure there is no other pet in the store? " When choices were made quickly and seemingly without much pretence behaviour, the examiner questioned: "What is this? . . . you wanted to choose a pet? "

Data Analysis.

The whole session was tape-recorded and transcribed verbatim. The complete protocols consisted of transcripts and the written records of the examiner. The data were analysed for type, incidence, and order of substitute selection, and also for degree of make-believe involvement as indicated by appropriate role enactment. The latter refers to the adoption of concrete role behaviours which are appropriate to the demands of the situation, for example, cuddling and feeding the hungry baby. Two independent scorers who were unfamiliar with the children scored all the records. Percentage of agreement was as follows: Feeding the Hungry Baby—96 per cent; Going to the Beach—97 per cent; the Petshop—96 per cent.

Each child participated in the puzzle and pretence games and the individual's performance on each set of tasks could thus be compared. In both situations the procedures were highly suggestive, freely encouraging substitutions.

RESULTS

The overall results of study 2, the pretence games, indicate that the child's selections were neither random nor arbitrary but followed a clearcut rule. The first choices focused on suitable objects: in the case of the hungry baby—food; in the case of the beach—the doll and other animate objects; in the case of the petshop—the kitten and other pets. When all the suitable objects had been exhausted, the child could be induced to make further selections, usually of objects whose characteristics were relatively neutral, i.e., indistinct in form and function. When only incongruous objects were available, the child's selections sharply diminished and refusals increased in frequency. This trend was clearly observable for all age groups and held true for the three pretence situations. Table 2 offers an example of the general trend for object selections. It presents the incidence of object selection for all subjects in the ' Feed the Hungry Baby ' game.

Task Effects.

While all three tasks were defined as pretence games, they differed, nevertheless, in terms of familiarity to the child and the power to elicit make-believe and role enactment. This was largely a function of the stimulus display and the specifications for action. Thus, the Hungry Baby game centred on a concrete baby doll and the instructions emphasised the need to feed the baby, a credible and well-known situation. In the Beach game the referent was merely a two-dimensional picture which represented a beach scene, while the invitation to bring a friend along failed to specify a particular pretence action and thus called for less direct involvement on the part of the player. In the Petshop the referent was evoked only verbally since a direct representation of a store was lacking, and the instruction to select a kitten also failed to specify precise pretence actions. The data reflect these differences in the structure of the game: pretence involvement as measured by appropriate role enactment was highest for the Hungry Baby game (mean score—4·38), and quite low for the Beach game (mean score—0·62) and the Petshop (0·38). A comparison of these scores (sign test) indicates that the differences between the pretence role enactment scores for the Hungry Baby and the Beach game are statistically significant ($P < \cdot 001$); the differences between the Hungry Baby and the Petshop are also statistically significant ($P < \cdot 001$). The small difference in scores obtained on the Petshop and Beach games is not statistically significant.

Task differences were also apparent in the child's willingness to go beyond the selection of suitable objects. In the Hungry Baby game only eight out of 60 children selected all possible substitutes, thus completing the whole series. In the Petshop game 11 children accepted all substitute objects offered in the array, and in the Beach game 24 children completed the whole series and accepted all of the available objects. Thus, an inverse relationship can be observed between pretence role enactment and substitute selections. The convincing Hungry Baby pretence situation yielded a higher mean role enactment score and a lower percentage of incongruous object choices. The less involving and more remote pretence situations of Petshop and Beach yielded a lower mean role enactment score and a higher incidence of incongruous selections. Sex differences were not recorded.

TABLE 2

DISTRIBUTION OF OBJECT SELECTION IN PERCENTAGES FOR 60 PRE-SCHOOL CHILDREN IN THE HUNGRY BABY GAME.

Bottle Small	*Bottle Medium*	*Bottle Large*	*Canned Food*	*Picture Food*	*Cardboard Roll*	*Play-dough Blob*	*Block*	*Candle Small*	*Candle Large*	*Pencil*	*Toy Animal*	*Toy Car*	*Hairbrush*
93	87	69	84	52	49	53	58	29	33	33	27	31	20

Age Effects.

While the order of substitute selections was comparable for all three age groups, on two tasks—the Petshop and the Beach games—the younger children (ages 2·8—3·7) made more incongruous substitute selections than the older ones. The younger children also showed a higher degree of role enactment on a single task, the Hungry Baby game. The mean role enactment score on this game for Group 1 (ages 2·8—3·7) was 5·5 while for groups 2 and 3 (ages 3·8—5·8) the mean score was 3·9.

Qualitative Observations.

The protocol material highlights the child's understanding of the substitute object and his or her approach to the task. When only incongruous objects were available children's responses extended from hesitation, evasiveness, or outright rejection to amusement, laughter and various forms of symbolic transformations.

The approach of individual children to the substitution task varied considerably. Some children engaged in various avoidance techniques designed to deny the need for an incongruous substitute object choice. Examples are statements to the effect that " the baby is asleep," " the baby is not hungry," " the baby is not crying," "she is getting too fat," "she is not going to eat now," "he is tired, he doesn't want to go to the beach," " I want to play at home, not on the beach."

Other solutions to the problem presented by the incongruous object array involved the use of mediators or reaffirmation of the real identity of the object. Thus, for example, several children used incongruous objects as a vehicle for the supply of the needed food: " I'll use the car for shopping," " I don't have any more (food)—I'll buy some at the store," " the rabbit will give the baby food." In the case of a reaffirmation of the dual nature of the pretence object, children remarked: " This is Ellen, but it is really a string," " it's really playdough," " she ate the brush," " this is Bennett (the boot)—he is heavy," " do you know it is poison? " (as she feeds the pencil to the baby), " she ate the car " (rabbit, or candle). These were efforts to resolve the discrepancy and eliminate the conflict; the child yielded to the suggestion but simultaneously reaffirmed the original meaning of the object. In a similar vein compromises were struck by hyphenating terms as the following examples illustrate: car-food, candle-food, rabbit-food, cube-sugar, a carrot-doggie, a playdough-pet.

Attempts to transform the substitute object in line with some of its own attributes , for example, colour or shape, can be seen in those cases where the yellow cube was identified as ' cheese ' or as an ' orange ' and the playdough blob became a ' meatball ' or a ' potato.'

Truly imaginative transformations relied heavily on verbal statements, for example, walking the substitute object toward the beach scene, while describing swimming, ball playing in the bright sunshine, and the building of sandcastles, accompanying the descriptions with appropriate gestures.

Pretence was frequently maintained by purely verbal means without gestural role enactment as the following examples indicate: " Who is going to be my friend? It's going to be him or her (points to giraffe and doll) . . . I think I'd better take the giraffe "; " I can put apple sauce in the large bottle . . . this is a strange bottle, good for apple sauce "; " here is a can of soup " (as he picks the cardboard roll); " this is a piece of butter on its bread " (points to cube); " she likes carrots " (hands the pencil to the examiner).

COMPARISON OF PUZZLE AND PRETENCE GAMES

The experimental situation for the two sets of games differed markedly. The puzzle task specified completion in terms of form, size, and meaning of a part, and seemed to demand a rather rigorous fit between the insert and the cut-out area. The pretence game, by definition, freely encouraged substitutions at the expense of form, size and meaning of the original object.

Comparison of the two sets of tasks yielded two major findings:

(*a*) A common rule determined the responses to both puzzle and pretence games, a rule which specified the order or sequence of item selection. On both tasks selection of substitute items proceeded from optimal choices to neutral ones, and trailed off with incongruous substitutes. The general trend was to reject incongruous items more frequently than suitable or neutral ones. Although our youngsters could be induced to accept less than optimal substitutes, eventually some and, at times, even all incongruous substitutes were rejected.

(*b*) The two tasks differed significantly in terms of the incidence of acceptable substitutions, i.e., the extent of substitute rejections. The pretence situation elicited a higher incidence of object substitutions (Wilcoxon Matched-Pairs Signed-Ranks Test). The differences between pretence and puzzle substitution scores are statistically significant ($P < \cdot 01$) for groups 2 and 3, ages 3·8—5·8. For group 1, the youngest children, ages 2·8—3·7, the differences are in the right direction just missing the level of statistical significance ($P < \cdot 07$).

DISCUSSION AND CONCLUSION

The results of this study confirm Sliosberg's original contention that pre-school children discriminate between pretence and realistic games. The evidence suggests that both play situations are rule-governed (though the ground-rules for each one differ) and obey certain limits. The extent to which children transform the object to fit the needs of the task is a function of the nature of the task (puzzle or pretence play) and the characteristics of the object. Even in the case where the task encourages transformations, there are apparently limits to the extent of object substitution. This is even more impressive if we consider the highly suggestive context provided by the experimenter and his positive reinforcement of the child's substitute selections.

The data also suggest that a child's perception of a task can undergo a playful transformation from requiring strict adherence to a formal rule of fit to a more humorous and symbolic one. This transformation is particularly evident in the puzzle situation where, for example, the child starts out with a rigid requirement to insert a piece which will fit the cut-out and ends up with such playful transformations as ' cat-man ' and ' halloween-mask.'

Our results have demonstrated the orderly character of the substitutions and thus appear to challenge Piaget's account of the presumably ' lawless ' substitutions of pretence play. Piaget, as will be recalled, stresses the undisciplined character of the young child's substitutions which supposedly follow the global principle that anything can stand for anything. Piaget perceives the spontaneous transformational activity typical of symbolic play as an expression of undisciplined thought, a position which in the light of our findings requires an adjustment.

Finally, we have to consider the interesting fact that our youngest children (2·8—3·7) made more substitutions on two of the three pretence games (Beach and Petshop) and on the three puzzles. This finding raises the following questions: (*a*) Is the greater number of substitute selections evidence of a lack of perceptual discrimination? (*b*) Is it evidence of erratic choices and undisciplined thinking? (*c*) Do the facts imply greater suggestibility and compliance?

The first alternative seems untenable in the light of the statements which indicated a clear recognition of the absurdity of the choices. The second alternative seems unlikely since the sequence of the selection of substitute items was orderly rather than arbitrary, and the age effect was only observed on two out of three pretence games. In the case of the Hungry Baby, the game which elicited the greatest degree of involvement on the part of the child, no such age effect was apparent. The third alternative suggests that the younger children were more susceptible to the influence of the adult who seemed to press for a continuation of the game, i.e., for further substitute selections. Considering the fact that our youngsters resisted this pressure on the one task which was most familiar and also elicited the greatest amount of role involvement, while yielding on the two games which elicited least involvement, one might venture the guess that the younger children recognised the rules of the game as played by the examiner and complied with his request on the games which were more remote and thus involved less conflict. On the Hungry Baby game, by contrast, familiarity and affective involvement imposed greater constraints on the selections. The notion of greater compliance with the adult's rules of the game might also underlie the younger children's somewhat larger number of substitute selections on the puzzle games. These are, however, *ad hoc* explanations which require further systematic study.

REFERENCES

EL'KONIN, D. (1966). Symbolics and its function in the play of children. *Soviet Educ.*, 8, 35-41.

FEITELSON, D., and ROSS, G. S. (1973). The neglected factor—play. *Human Dev.* 16, 202-223.

PIAGET, J. (1962). *Play, Dreams and Imitation.* New York: Norton.

SLIOSBERG, S. (1934). Zur Dynamiek des Ersatzes. *Psychologische Forschung*, 19, 122-181.

SMILANSKY, S. (1968). *The Effects of Sociodramatic Play on Disadvantaged Preschool Children.* New York: Wiley.

SUTTON-SMITH, B. (1967). The role of play in cognitive development. *Young Children*, 6, 364-369.

(*Manuscript received 6th September, 1976*)

Br. J. educ. Pychol., 47, 187-192, 1977

THE EFFECTS OF CHRONOLOGICAL AGE ON THE RELATIONSHIP OF ACADEMIC ACHIEVEMENT WITH EXTRAVERSION AND NEUROTICISM: A FOLLOW-UP STUDY

BY G. M. SEDDON
(School of Chemical Sciences, University of East Anglia).

SUMMARY. A sample of 741 students worked through a self-instructional chemistry programme after completing tests of extraversion, neuroticism, verbal intelligence, non-verbal intelligence and background knowledge in chemistry. The existence of significant interactions involving extraversion and neuroticism with the measures of chronological age, intelligence and background knowledge in determining post-test performance was then investigated. The results showed that none of the interactions involving neuroticism was significant, and indicated that the significant interaction involving extraversion and chronological age was not due to the parallel effects of verbal intelligence, non-verbal intelligence or background knowledge.

INTRODUCTION

IN recent years there has been considerable interest in the suggestion that the relationship of academic achievement with both extraversion and neuroticism may vary with chronological age (Entwistle, 1972; Anthony, 1973). The bulk of the evidence quoted in favour of such interactions is based upon a comparison of the results of different investigations in which different tests of academic achievement have been administered to students of different ages (Entwistle, 1972). However, on moving from one investigation to the next, changes in chronological age are invariably confounded with concomitant changes in the nature of the tests. Furthermore, both the magnitude and direction of the observed effects, which had previously been attributed to interactions involving chronological age, could equally well have been caused by interactions involving the nature of the test (Seddon, 1975). Thus, these studies do not provide conclusive evidence for interactions involving chronological age.

There appear to have been only two relevant published studies which have been specifically designed to avoid this particular confounding of variables (Leith and Davis, 1972; Seddon, 1975). The study by Leith and Davis was concerned with investigating the relationship between the amount learned from a self-instructional programme by 12-and 13-year-olds, and demonstrated that the regression line of the amount learned on neuroticism followed the shape of an inverted-U. Among older students the shape of the regression curve was modified by the maximum of the curve moving towards higher levels of neuroticism. The study by Seddon concerned students in the 15 to 19+ age range and involved the administration of a battery of intelligence and achievement tests to 741 chemistry students drawn from GCE O- and A-level classes in secondary schools as well as first-year university classes. The results showed that in the case of achievement tests there were interactions between extraversion and chronological age, which operated so as to increase the partial relationship of achievement with extraversion as chronological age increased. However, there were no such interactions involving neuroticism. Neither were there interactions involving either extraversion or neuroticism in determining the performance on the intelligence tests.

Thus, both studies indicate that the suggested interactions involving chronological age do appear to operate in the case of achievement. However, in both studies it is important to recognise that as students progress from the lower to the higher levels of the educational system there are concomitant changes in other variables, which may be of importance in determining the level of a student's subsequent academic achievement. In particular, there is continual intellectual growth during adolescence (Bayley, 1970), paralleled by increases in general background knowledge. In addition, the students who took part in Seddon's (1975) study were at different stages in a sequence of selection procedures, which considerably reduce the number of students transferring from the pre- O-level courses at the age of around 16 years, through the A-level courses to the university level courses. As a result the students at the different ages covered distinctly different ranges as regards general intelligence and background knowledge. Hence, while the two studies demonstrate that there are interactions which relate to chronological age, it is not clear whether the interactions refer to the intrinsic effects of age, such as processes of maturation, or to changes in other variables, which are themselves incidentally related to age, for example, the effects of learning. As Entwistle and Welsh (1969) showed that at the age of 14 the relationship of achievement with extraversion changes from being negative with boys of high ability, to being negative with boys of low ability, there is even more likelihood that the interactions could be attributed to the effects of general intelligence and previous knowledge. Hence the purpose of the present investigation is to follow up the previous study (Seddon, 1975) in order to examine this possibility.

METHOD

The basic plan was to administer tests of extraversion (E), neuroticism (N), verbal intelligence (VI), non-verbal intelligence (NVI) and background knowledge in chemistry (BKC) to a sample of students covering a substantial range of chronological ages (CA), and then to investigate the existence of the various interactions of interest in determining a student's subsequent achievement (A) after learning from a self-instructional programme, which was common to all the students. In particular, it was important to discover whether the relationship of achievement with each of extraversion and neuroticism was dependent upon interactions with chronological age, with each form of intelligence or with general background knowledge.

If there were interactions due to the intrinsic effects of chronological age, these interactions would be observed using the raw measures of chronological age. If, in addition, the results showed that there were no interactions involving intelligence or background knowledge, it would be reasonable to conclude that the observed interactions involving chronological age were due to some intrinsic effects of chronological age unconnected with parallel changes in intelligence or background knowledge. It would also be reasonable to draw a similar conclusion if the interactions involving intelligence and background knowledge were significant but with the effects of increasing intelligence and increasing background operating in the opposite direction to that obtained from increasing chronological age. For example, such a situation would arise if the relationship of achievement with extraversion increased on increasing chronological age, but decreased on increasing the intelligence scores. On the other hand, if the significant interactions involving intelligence or background knowledge operated in the same direction as the interactions involving chronological age, it would be necessary to repeat the statistical analysis using measures of achievement, chronological age, intelligence, and background knowledge, which had been suitably corrected in order to separate out the effects of chronological age from those of intelligence and background knowledge.

Sample.

The 741 students who took part in the present experiment were the same sample used in the previous study (Seddon, 1975). They were all students of chemistry in either secondary or tertiary institutions. The 25 secondary schools included comprehensive, grammar and public schools, drawn mainly from the Midlands. Each school was preparing the students for either O- or A-level chemistry examinations. From the higher education sector two colleges of education and three universities allowed all the first-year chemistry students, who were potential candidates for honours degrees, to participate. The numbers and distributions of the students among the different age groups and different types of institution are summarised in Table 1.

TABLE 1

DISTRIBUTION OF THE SUBJECTS AMONG THE DIFFERENT AGE GROUPS AND TYPES OF ACADEMIC INSTITUTION.

Age group (years)	Schools	Colleges of Education	Universities
15	97	0	0
16	247	0	0
17	178	0	12
18	11	0	141
19+	0	21	34
All groups	533	21	187

The tests.

The tests of extraversion and neuroticism were those contained in the Eysenck Personality Inventory (Eysenck and Eysenck, 1964). Verbal and non-verbal intelligence were measured by the AH5 (Heim, 1956), while the test of background knowledge in chemistry was 'High School Chemistry' form 1965 (Part II), published by the American Chemical Society—National Science Teachers Co-operative Examinations Committee.

The self-instructional programme.

A programme entitled " Chemical Structures and Physical Properties " was specially prepared by the present author. Its objectives were those which had been tested by the test of specific background knowledge used in the previous study (Seddon, 1975). The students were expected to identify graphs showing the correct trends in the magnitudes of various physical properties of binary compounds in certain parts of the Periodic Table of the Elements, and, secondly, to identify diagrams showing the three-dimensional molecular or ionic structures which could be adopted by these substances in the solid state. Naturally, the programme assumed some initial knowledge on the part of the student. However, the programme began at a level which was considered suitable for the youngest students, i.e., those who, in general, would be expected to have the least amount of relevant background knowledge—and every attempt was made to include all the concepts and principles which these students would require. Hence, depending upon the effectiveness of this teaching, those students who might be at a disadvantage in terms of initial knowledge were given an opportunity to catch up.

The programme underwent several trials in institutions similar to those participating in the final experimental run. After each trial it was revised in the light of the results from these trials. The programme was finally produced as a booklet comprising five sections, each section containing 40 frames, which could be read by the students in 30 to 40 minutes.

The pre- and post-tests were one and the same. The test of specific background knowledge in chemistry, described in the previous study, was used in both cases.

Experimental procedure.

Since the present and previous studies had been planned to take place consecutively in one complete exercise, the measures of chronological age, together with the scores on extraversion, neuroticism, intelligence, background knowledge and pre-tests, were those which had been used in the previous study. Thus, the intelligence and personality tests were administered some weeks in advance of the other tests and materials, and all the remaining tests and materials were administered in nine one-hour periods spread over a period of three weeks. During this interval of three weeks none of the institutions taught their students subject matter which was directly related to the content of the self-instructional programme. As it proved to be impossible for all the institutions to use the same period of three weeks, they all took part at various times over an interval of three months. The classroom periods during which the four tests and five sections of the self-instructional programme were each administered, were one hour in length, and were always supervised—usually by the class teacher.

In most of the institutions it was administratively impossible to allocate some of the students to a control group in which the students did not read the self-instructional programme at all. Thus, all the students in a particular class read the self-instructional programme.

RESULTS

The mean scores for the pre- and post-tests were respectively 20·6 and 31·3. The corresponding standard deviations were 10·6 and 11·0. The difference between the mean scores on the pre-and post-tests is significant at the ·01 level on a one-tail test (t=33, df=740).

A suitable method of investigating the existence of the interactions is multiple regression analysis. The various equations, which were constructed, are modelled on equation (1) where X_D is a general term to represent measures of chronological age, verbal intelligence, non-verbal intelligence and general background knowledge.

$$\hat{X}_A = b_1.X_D + b_2.X_E + b_3.X_N + b_4.X_E^2 + b_5.X_N^2 + b_6.X_D.X_E + b_7.X_D.X_N + \text{const} \quad (1)$$

The interactions of extraversion and neuroticism with each of these variables are represented by the cross-product terms $X_D.\ X_E$ and $X_D.\ X_E$ respectively.

The results relating to each of the instances of D are shown in Table 2 which presents a whole series of multiple regression equations gradually building up to the final equation by the addition of a new term on moving from one equation to the next.

TABLE 2

RESULTS OF STEPWISE MULTIPLE REGRESSION ANALYSIS FOR THE DIFFERENT EQUATIONS RELATING TO CHRONOLOGICAL AGE, VERBAL INTELLIGENCE, NON-VERBAL INTELLIGENCE, AND BACKGROUND KNOWLEDGE IN CHEMISTRY.

CHRONOLOGICAL AGE.*

Variable	beta-weights ($\times 10^2$)							
CA	22	21	21	21	22	01	27	–16
E		–18	–19	–08	–07	–27	–09	29
N			–06	–06	07	06	10	01
E^2				–11	–12	–05	–11	–06
N^2					–14	–13	–14	–14
CA × E						26		32
CA × N							–07	13
R($\times 10^2$)	22	29	30	30	30	32	30	32
P<	·01	·01				·01		

*Ages are measured to the nearest quarter of a year.

NON-VERBAL INTELLIGENCE.

Variable	beta-weights ($\times 10^2$)							
NVI	22	21	21	21	21	57	16	54
E		–19	–20	–12	–12	37	–12	36
N			–04	–04	06	08	–01	03
E^2				–08	–08	–10	–08	–09
N^2					–10	–12	–11	–12
NVI × E						–59		–58
NVI × N							09	06
R($\times 10^2$)	22	29	29	29	29	31	29	31
P<	·01	·01				·01		

VERBAL INTELLIGENCE

Variable	beta-weights ($\times 10^2$)							
VI	33	32	32	32	32	53	49	72
E		–19	–19	–06	–06	21	–06	22
N			–03	–03	06	06	32	33
E^2				–13	–13	–14	–14	–15
N^2					–09	–09	–08	–08
VI × E						–34		–35
VI × N							–32	–33
R($\times 10^2$)	33	38	38	38	38	38	38	39
P<	·01	·01						

BACKGROUND KNOWLEDGE IN CHEMISTRY.

Variable	beta-weights ($\times 10^2$)							
BKC	61	59	59	59	59	72	69	82
E		–08	–08	–01	00	14	–01	15
N			00	00	14	14	24	24
E^2				–08	–08	–13	–08	–13
N^2					–15	–15	–15	–14
BKC × E						–14		–15
BKC × N							–13	–14
R($\times 10^2$)	61	62	62	62	62	62	62	62
P<	·01	·01						

In this matrix format the beta-weights and values of multiple-R for any one equation in this stepwise procedure are presented on one column, and the independent variables themselves are given in the order in which they were introduced, down the left-hand column of the matrix. The significance of each term is assessed by testing whether or not the predictability is significantly improved after adding that particular term (McNemar, 1969, p. 321). In cases where terms are significant at the 1 per cent level, this fact is indicated at the foot of the appropriate column. In investigating the significance of each of the two interaction terms, the significance test is carried out only when the other term had already been introduced into the equation. This precaution was taken in order to ensure that the significance of one or other of these two terms was not missed, due to the possible existence of mutually suppressing effects.

DISCUSSION

The significant difference between the mean pre-and post-test scores is taken to indicate that learning had taken place. Since the measures of chronological age, intelligence and background knowledge were obtained before the students learned, it is reasonable to regard them as measures which may determine a student's performance on the post-test after learning.

As regards the interactions, none of those involving neuroticism is significant. In the case of extraversion two interactions are significant, one involving non-verbal intelligence, and the other involving chronological age. However, as chronological age and non-verbal intelligence both increase, the interaction between extraversion and non-verbal intelligence operates in the opposite direction to that of the interaction between extraversion and chronological age. Since there was no significant interaction which operated in the same direction as those involving chronological age, the overall pattern of results may be interpreted, as they are, without doing any covariance adjustments to remove the effects of chronological age from those of both forms of intelligence and background knowledge.

The absence of an interaction involving neuroticism and chronological age is in line with the result of the previous study (Seddon, 1975), but in contrast to the result obtained by Leith and Davis (1972) with 12-and 13-year-olds. Once again it may be that this difference is in some way related to the different age ranges of the students in the different experiments.

The existence of the interaction between extraversion and age also corresponds to the result of the previous study. Furthermore, the fact that the parallel interactions with intelligence or background knowledge were either not significant or operated in the reverse direction to that of the interaction with chronological age, indicates that the effects of chronological age are quite distinct from those of intelligence or background knowledge. In this respect it is also significant that, when all the terms have been introduced into each multiple regression equation, the partial relationship between achievement and age is negative, whereas in the cases of each form of intelligence and background knowledge it is positive.

In conclusion the present investigation appears to have eliminated the possibility that the significant interaction relating to chronological age could have been due to the concurrent effects of verbal intelligence, non-verbal intelligence or background knowledge. The interactions are due either to the intrinsic effects of chronological age, such as maturational processes, or to the concurrent effects of other variables which have not been investigated in this study.

REFERENCES

ANTHONY, W. S. (1973). The development of extraversion, of ability and of the relation between them. *Br. J. educ. Psychol.*, 43, 223-227.

BAYLEY, N. (1970). Development of mental abilities. In MUSSEN, P. H. (Ed.), *Carmichael's Manual of Child Psychology* (3rd edit.), Vol. 1. New York: Wiley.

ENTWISTLE, N. J. (1972). Personality and academic attainment. *Br. J. educ. Psychol.*, 42, 137-151.

ENTWISTLE, N. J., and WELSH, J. (1969). Correlates of school attainment at different ability levels. *Br. J. educ. Psychol.*, 39, 57-63.

EYSENCK, H. J., and EYSENCK, S. B. G. (1964). *Manual of the Eysenck Personality Inventory.* London: University of London Press.

HEIM, A. W. (1956). *AH5 Group Test of High-grade Intelligence.* Slough : National Foundation for Educational Research.

LEITH, G. O. M., and DAVIS, T. N. (1972). Age changes in the relationship between neuroticism and achievement. *Res. in Educ.*, 8, 61-70.

MCNEMAR, Q. (1969). *Psychological Statistics* (4th edit.). New York: Wiley.

SEDDON, G. M. (1975). The effects of chronological age on the relationship of intelligence and academic achievement with extraversion and neuroticism. *Br. J. Psychol.*, 66, 493-500.

(Manuscript received 4th June, 1976)

RESEARCH NOTES

Br. J. educ. Pychol., 47, 193-196, 1977

THE DEVELOPMENT OF EXTRAVERSION AND ABILITY : AN ANALYSIS OF RUSHTON'S LONGITUDINAL DATA

By W. S. ANTHONY
(*School of Education, University of Leicester*)

Summary. Correlations between extraversion and ability typically change from positive to negative at about 13–14 years. A method for analysing this change is developed and applied to Rushton's (1969) data from his longitudinal study of 266 children of above-average intelligence, whom he tested at 10–11 years and again at 15–16 years of age. Correlations between extraversion and relative decrease in English and in mathematics are found and attributed to an association between introversion and study. A correlation between intelligence and relative decrease in extraversion is found and attributed to a persistent earliness *versus* lateness of development. An appendix indicates a statistical proof and provides convenient formulae.

INTRODUCTION

Up to the age of around 13–14 years, extraversion scores and ability scores are usually positively correlated, whereas after that age they are usually negatively correlated (Eysenck and Cookson, 1969; Elliott, 1972; Entwistle, 1972). Here, the term ' extraversion ' includes Eysenckian and Cattellian extraversion scores, and ' ability ' includes intelligence and educational attainment scores; this inclusiveness is convenient for stating the broad generalisation even though distinctions are drawn by reviewers.

Now, a change in correlation is a statistical change which must be constituted out of changes in individual persons. A crowd cannot move east if everybody in the crowd is still or moves west; similarly, though in a rather more complex sense, a correlation cannot change unless there are appropriate changes in the relative positions of the individuals. Is it, then, crudely speaking, that the more able children become introverted while the less able children become extraverted (which would account for the observed change in correlation), or, alternatively, is it that the extraverts fall behind in the development of ability while the introverts make faster progress (which would also account for the change in correlation)? Or is there a combination of both kinds of transformation? Also, how should these hypotheses be reformulated to avoid the crude dichotomies of extravert-introvert and ' more able ' *versus* ' less able ' ?

The purpose of this paper is to report an investigation of these matters by reanalysing data collected by Rushton (1969) in his longitudinal study of 266 children, of above-average intelligence, tested at 10–11 years and again at 15–16 years. Interest in this problem was stimulated by Rushton's earlier paper (Rushton, 1966) and by Wilson (1972).

METHOD AND RESULTS

The task is to analyse the change in extraversion-ability correlations in such a way as to reveal the nature of the changes in the individual children. The crude idea that the more able children become introverted while less able children become extraverted, may be reformulated as the idea that ability is correlated with decrease in extraversion. The second alternative, mentioned earlier, may similarly be reformulated in terms of extraversion being correlated with decrease in ability. But how are these hypothetical correlations to be calculated, and how are they related to the observed change in extraversion-ability correlations?

The following statistical method was devised. First, it is convenient to suppose that all test scores have been converted to standard z scores. Let the symbols A1, A2, E1 and E2 represent a child's ability score at first testing (age 10–11 years), his ability score at second testing (15–16 years), his extraversion score at first testing, and his extraversion score at second testing, respectively. Each of the four sets of scores is conceived to have been separately standardised (mean=O; SD=1). 'Ability' is then defined in terms of average scores, $A=\frac{1}{2}(A1+A2)$ for any individual child, while 'decrease in extraversion' is defined as $E1-E2$. Similarly, 'extraversion' is defined as $E=\frac{1}{2}(E1+E2)$, and 'decrease in ability' is defined as $A1-A2$.

The following equation (see Appendix for derivation) shows the analysis of the change in the extraversion-ability correlation:

$$ {}^{r}A1E1 - {}^{r}A2E2 = [{}^{r}A, E1-E2]({}^{\sigma}A\,{}^{\sigma}E1-E2) + [{}^{r}E, A1-A2]({}^{\sigma}E\,{}^{\sigma}A1-A2)\cdot $$

The left-hand side is simply the change which is to be analysed. On the right-handside of the equation, the first correlation represents the idea that more able children become introverts while the less able children become extraverts. The second correlation in square brackets is between extraversion and decrease in ability, representing the idea that extraverts decrease in ability while introverts increase in ability. However, 'decrease in ability' means decrease in standard-score ability, that is, decrease relative to the other children's development, not an absolute decline. The definition of a child's ability as $A=\frac{1}{2}(A1+A2)$ perhaps needs clarification; this average may be regarded as a factor common to the two ability scores, that is, it measures what *endures* between the two testings of ability, whereas $A1-A2$ measure what *changes* between the two testings. Similarly, of course, with respect to 'E' and '$E1-E2$.' Thus, the correlations in square brackets represent the two kinds of possible transformation in children which were envisaged in the introduction. The standard deviations in parentheses are of little interest.

TABLE 1

MODEL EQUATION APPLIED TO RUSHTON'S LONGITUDINAL DATA.

Model:

$$ {}^{r}A1E1 - {}^{r}A2E2 = [{}^{r}A, E1\text{-}E2]({}^{\sigma}A\,{}^{\sigma}E1-E2) + [{}^{r}E, A1-A2]({}^{\sigma}E\,{}^{\sigma}A1-A2) $$

Extraversion and Intelligence:

·142*—(—·088)=[·175**] (1·117)+[·043] (·800)

Extraversion and Mathematical Ability:

·128*—(—·136*)=[·067] (1·164)+[·262**] (·709)

Extraversion and Ability in English:

·226**—(—·239**)=[·241**] (·882)+[·226**] (1·117)

NOTES: Correlations marked * and ** are significant at the ·05 and ·01 levels, respectively. Correlations on the left-hand side are Rushton's own; correlations and other terms on the right-hand side have been calculated by the present writer from Rushton's correlation matrix. Rushton normalised scores before correlating them.

Table 1 shows the application of the model equation to Rushton's (1969) data. The extraversion scores were obtained from Cattell's Children's Personality Questionnaire (CPQ) at 10–11 years, and from Cattell's High School Personality Questionnaire (HSPQ) at 15–16 years. For the purpose of the present analysis, Moray House Verbal Reasoning Quotient at 10–11 years, and Factor B (Intelligence) on HSPQ at 15–16 years, were taken as measures of intelligence. Rushton deliberately selected his sample to be above average in intelligence: the mean Verbal Reasoning Quotient was 111·8. Table 1 shows that the extraversion-intelligence correlation changed from +·142 to —·088. The Table also shows that the correlations between intelligence and decrease in extraversion (+·175) is significant, whereas the correlation between extraversion and decrease in intelligence (+·043) is not.

Table 1 also presents the application of the method to the correlations obtained by Rushton between extraversion and other measures of ability. Arithmetic Quotient (at 10–11 years) and the GCE O-level Mathematics scores were considered to be plausible A1 and A2 scores in mathematical ability after conversion to standard scores. The English Quotient (at 10–11 years) and the GCE O-level English Language scores were likewise taken as scores in ability in English. (The correlation between English Quotient and GCE English was slightly negative, — ·105; Thus, it might seem improper to average the standard scores to make a single score in English for each person ; however, a correlation of — ·105 implies substantial correlations (+ ·67) between each separate score-variable and the common factor formed by the average.)

Table 1 shows that the results in respect of mathematics and English differ in some ways from those relating to intelligence: the change in extraversion-ability correlations is again found to be from positive to negative, but the correlations between extraversion and decrease in these educational attainments are significant, whereas the correlation between extraversion and decrease in intelligence was not.

CONCLUSIONS

The following two conclusions are based on small, highly significant, correlations. (1) The more intelligent children and those able in English, tended to become relatively more introverted than the less intelligent and the less able in English. (2) The more extraverted children tended to become less able in English and mathematics than the introverted children. Previously the change in extraversion-ability correlations could be hypothetically attributed to *either* of the two kinds of transformation envisaged in the introduction, whereas it is concluded here that *both* kinds of transformation are significantly involved.

The results suggest that the prevalence of one or other kind of transformation depends on the kind of ability which is tested. The correlation between introversion and increasing relative success in academic examinations is plausible, since such success is presumably facilitated by private study which is an introverted (withdrawn, cautious, sober) type of behaviour. A similar conclusion arises from the work of Banks and Finlayson (1973), who studied secondary-school boys of above-average ability, and found that Eysenckian introversion was linked with intellectual curiosity, commitment to homework, and with increase in academic success relative to predictions from scores at the end of primary school.

Turning now to the correlation between intelligence and decreasing extraversion, this may be attributed to an earliness *versus* lateness of development in the 10–16 age-period. A consistently and persistently early developer will be more intelligent than his later-developing age-mates but he will decline early in extraversion, while *their* extraversion is still increasing (Anthony, 1973).

In classical test theory, observed scores are unreliable and this affects the correlations obtained. Corrections for attenuation may be applied to the correlations; however, Lord and Novick (1968) show that correlations tend to be over-corrected, and Lumsden (1976), in a vigorous review, declares that the correction should never be used. Although the results of such corrections might (if warnings are disregarded) be used to analyse estimated *theoretical* changes in extraversion-ability correlations, they would still have no apparent place in the analysis of *observed* changes.

The results reported here are derived solely from Rushton's data and may thus depend on the particular circumstances of his research, such as the nature of the children and of the test instruments. But other relevant longitudinal data are known to exist, where different tests have been used with different kinds of children, and it is hoped that these data may be subjected to a similar analysis.

ACKNOWLEDGMENT—Dr. Rushton's consent to this use of his data is gratefully acknowledged.

REFERENCES

ANTHONY, W. S. (1973). The development of extraversion, of ability, and of the relation between them. *Br. J. educ. Psychol.*, 43, 223-227.

BANKS, O., and FINLAYSON, D. (1973). *Success and Failure in the Secondary School.* London: Methuen.

ELLIOTT, C. D. (1972). Personality factors and scholastic attainment. *Br. J. educ. Psychol.*, 42, 23-32.

ENTWISTLE, N. J. (1972). Personality and academic attainment. *Br. J. educ. Psychol.*, 42, 137-151.

EYSENCK, H. J., and COOKSON, D. (1969). Personality in primary school children: 1—Ability and Achievement. *Br. J. educ. Psychol.*, 39, 109-122.

LORD, F. M., and NOVICK, M. R. (1968). *Statistical Theories of Mental Test Scores.* Reading, Mass.: Addison-Wesley.

LUMSDEN, J. (1976). Test theory. *Ann. Rev. Psychol.*, 27, 251-280.

RUSHTON, J. (1966). The relationship between personality characteristics and scholastic success in 11-year-old children. *Br. J. educ. Psychol.*, 36, 178-184.

RUSHTON, J. (1969). *A Longitudinal Study of the Relationship Between Some Personality Variables and Some Measures of Academic Attainment* Unpublished Ph.D. thesis, University of Manchester.

WILSON, J. A. (1972). Personality and attainment in the primary school. II—Personality structure and attainment of 10-year-olds. *Res. in Educ.*, 7, 1-10.

(*Manuscript received 9th June*, 1976)

APPENDIX

The equation used in the text of this paper can be proved from the following considerations: First, by ordinary algebra, we find that

$$(A_1+A_2)\,(E_1-E_2)+(E_1+E_2)\,(A_1-A_2)=2A_1E_1-2A_2E_2$$

If we reverse sides, divide by 2, sum over N individuals, and divide by N, we find

$$\frac{\Sigma A_1E_1}{N}-\frac{\Sigma A_2E_2}{N}=\frac{\Sigma\frac{1}{2}(A_1+A_2)\,(E_1-E_2)}{N}+\frac{\Sigma\frac{1}{2}(E_1+E_2)\,(A_1-A_2)}{N}$$

Now, bearing in mind that A_1, E_1, A_2, E_2, are all in *z* score form, it can be shown that each of the terms in the equation above is equal to the corresponding term in the equation below, i.e.,

$$A_1E_1-{}^{r}A_2E_2={}^{r}\tfrac{1}{2}(A_1+A_2),\,E_1-E_2\,{}^{\sigma}\tfrac{1}{2}(A_1+A_2)\,{}^{\sigma}E_1-E_2\;+$$

$${}^{r}\tfrac{1}{2}(E_1+E_2),\,A_1-A_2\,{}^{\sigma}\tfrac{1}{2}(E_1+E_2)\,{}^{\sigma}A_1-A_2$$

which is the equation used in the text. In applying the method it is convenient to obtain the required empirical values by the following formulae. Although X_1, Y_1, etc. must be conceived to be in *z* score form, the correlations in the right-hand sides of the formulae below are equivalent to the ordinary correlations prior to *z* score conversion, so the calculations can proceed from an ordinary correlation matrix.

$${}^{r}\tfrac{1}{2}(X_1+X_2),\,Y_1-Y_2=\frac{{}^{r}X_1Y_1-{}^{r}X_1Y_2+{}^{r}X_2Y_1-{}^{r}X_2Y_2}{{}^{\sigma}X_1+X_2\,{}^{\sigma}Y_1-Y_2}$$

$${}^{\sigma}X_1+X_2=\sqrt{2(1+{}^{r}X_1X_2)}\;;\quad{}^{\sigma}\tfrac{1}{2}(X_1+X_2)=\tfrac{1}{2}\,{}^{\sigma}X_1+X_2$$

$${}^{\sigma}X_1-X_2=\sqrt{2(1-{}^{r}X_1X_2)}$$

Br. J. educ. Psychol., 47, 197-198, 1977

THE RELIABILITY OF THE MATCHING FAMILIAR FIGURES TEST

By E. CAIRNS
(*Psychology Department, The New University of Ulster*)

SUMMARY. Despite the large amount of interest shown in Kagan's conceptual tempo dimension only limited evidence is available regarding the reliability of the principal test, the Matching Familiar Figures test. The present study examined the split-half reliability of the MFF in 50 9-year-old boys and 50 12-year-old boys. Response time reliabilities were adequate, but error score reliabilities were less than satisfactory.

INTRODUCTION

In 1964 Kagan and his co-workers claimed to have identified a cognitive style which distinguishes a reflective from an impulsive mode of information processing in children. The ' most sensitive index ' (Kagan, 1967) of this dimension is the Matching Familiar Figures test, a visual matching-to-sample task which provides two variables, time to first response and errors (a maximum of six per item are permitted).

Ten years after its inception conceptual tempo is now considered a well documented cognitive style, regarded as one of the few cognitive styles for which " reasonable numbers of properly designed investigations " have been carried out (Cashdan and Lee, 1971), included in a grand design for a " Conceptual Framework for a Multi-Factor Theory of Individuality " (Royce, 1973) and hailed as being of major significance for education (Rohwer, 1970). However, a critical omission from the literature to date has been any attempt to evaluate the reliability of the principal measure of conceptual tempo, the Matching Familiar Figures Test.

At present some limited evidence is available regarding the reliability of the MFF. Kagan (1965) retested 46 boys and 56 girls in grade one over a period of one year. The resulting correlations for response time were ·48 and ·50 for boys and girls, respectively (both $P<0{\cdot}01$), while for error scores they were ·25 (NS) for boys and ·51 ($P<0{\cdot}01$) for girls. Less convincing results were obtained in a later study (Yando and Kagan, 1968) in which MFF retest after 7–8 months provided reliability coefficients of ·70 ($P<0{\cdot}01$) and ·13 (NS) for 80 boys and 80 girls (grade one) on response time measures and ·24 ($P<0{\cdot}05$) and ·23 ($P<0{\cdot}05$) on error scores.

Siegelman (1969) reports on the reliability of the MFF over a shorter period of time and with older children but again the results are probably contaminated on several counts. She retested 65 grade four boys and after a five-week period, on a mechanised version of the MFF, in which the position of the stimuli were altered and their size reduced from that of the original version. The correlations she reports, both for response time (·73, $P<{\cdot}01$) and for errors (·43, $P<{\cdot}01$) were, therefore, probably affected by these changes and inflated by the fact that she did not retest 35 of her original subjects who failed to meet a double median-split criterion on the first test. The error score correlations were also further affected because on the second test a maximum of one error per trial only was permitted, instead of the usual six.

The best evidence on the test-retest reliability of the MFF (over a period of three weeks) is presented by Adams (1973), in a study involving 50 first-grade children (boys and girls). For response time the correlation was ·58 ($P<0{\cdot}01$, both sexes combined) while for error scores only those for girls are reported ·39 ($P<0{\cdot}01$). Adams concludes that the MFF is a reliable instrument " at least for research purposes."

It would appear that there is a lack of convincing evidence especially regarding the reliability of the MFF over short intervals and with older children. As the test is now being recommended for diagnostic use in education, not simply for research purposes, more information is required and the present study, therefore, examined the reliability of the MFF in older children using a split-half method.

METHOD

The sample consisted of 50 boys aged 9 years and 50 boys aged 11 years. Each child was seen individually for the administration of the two practice and 12 test items of the MFF: All tests were administered by the same person.

RESULTS

In order to carry out the examination of the internal consistency of the MFF, items were assigned to two sets according to two criteria—position of the item in the first or second half of the whole test and position of the target (the correct answer) for that item. The final division of the items resulted, thefore, in two six-item sets each containing one item at each of the six target positions and three items from the first and three from the second half of the test.

For the 9-year-old response times correlated ·94 ($P<0\cdot001$) and errors ·46 ($P<0\cdot01$) while for the 12-year-old group the corresponding correlations were ·95 ($P<0\cdot001$) and ·52 ($P<0\cdot001$). To estimate the reliability of the full test, that is the 12-item MFF, the Spearman-Brown formula was applied. This suggests that the 12-item MFF response time reliability coefficient is ·96 for 9-year-olds and ·97 for 11-year-olds and that the error score equivalents are ·63 and ·68 respectively.

DISCUSSION

The results of the present study confirm the impression created by existing studies that while MFF response time is a reasonably reliable measure, the reliability of error scores is less than satisfactory. This raises some serious implications for most conceptual tempo studies which have made use of error scores, particularly the wisdom of the inclusion of errors in the operational definition of reflective and impulsive performance (Kagan *et al.*, 1966). Further, the suggestion that MFF error scores are not a reliable measure may help to explain why investigators have met with difficulty in attempting to modify error scores, and why at least one study (Debus, 1970) has reported a control group increase in error scores on retest. Also this lack of evidence for the reliability of the 12-item MFF must call into question results obtained using 5-item (Drake, 1970) and 6-item (Meichenbaum and Goodman, 1971) versions. Obviously what is required is a more reliable version of the MFF: until such a version is available investigators should treat MFF error scores with caution.

REFERENCES

ADAMS, W. V. (1973). Strategy differences between reflective and impulsive children. *Child Dev.*, 43, 1076-1080.

CASHDAN, A., and LEE, V. (1971). *Learning Styles.* Milton Keynes: Open University Press.

DEBUS, R. J. (1970). Effects of brief observation of model behaviour on conceptual tempo of impulsive children. *Dev. Psychol.*, 20, 22-32.

DRAKE, D. M. (1970). Perceptual correlates of reflective and impulsive behaviour. *Dev. Psychol.*, 2, 202-214.

KAGAN, J. (1965). Reflection impulsivity and reading ability in primary grade children. *Child Dev.*, 36, 609-628.

KAGAN, J. (1967). Biological aspects of inhibition systems. *Amer. J. dis. Child.*, 114, 507-512.

KAGAN, J., PEARSON, L., and WELCH, L. (1966). Modifiability of an impulsive tempo. *J. educ. Psychol.*, 57, 359-365.

KAGAN, J., ROSMAN, B. L., DAY, D., ALBERT, J., and PHILLIPS, W. (1964). Information processing in the child: significance of analytic and reflective attitudes. *Psychol. Mono.*, 78, 1-37.

MEICHENBAUM, D., and GOODMAN, J. (1971). Training impulsive children to talk to themselves: a means of developing self-control. *J. Abnorm. Psychol.*, 77, 115-126.

ROHWER, W. D. (1970). Implication of cognitive development for education. In MUSSEN, P. H. (Ed.), *Carmichael's Manual of Child Psychology.* New York: Wiley.

ROYCE, J. R. (1973). The conceptual framework for a multi-factor theory of individuality. In ROYCE, J. R. (Ed.), *Multivariate Analysis and Psychological Theory.* London: Academic Press.

SIEGELMAN, E. (1969). Reflective and impulsive observing behaviour. *Child Dev.*, 40, 1213-1222.

YANDO, R. M., and KAGAN, J. (1968). The effect of teacher tempo on the child. *Child Dev.*, 39, 27-34.

(*Manuscript received* 4*th June*, 1976)

Br. J. educ. Psychol., 47, 199-202, 1977

TREATMENT OF ENCOPRESIS IN A CLASSROOM SETTING: A CASE STUDY

By ELIZABETH A. SCOTT
(*Educational Psychologist, Grampian Region, Aberdeen*)

SUMMARY. This study describes the procedure and results of a behaviour modification programme carried out in the classroom and aimed at eliminating encopresis in an 8-year-old boy. Initially, when the teacher assumed the role of therapist, the programme failed. A nursery nurse was then taught the use of differential reinforcement paired with a time-out contingency and within one week frequency of soiling decreased dramatically. This level of eliminative behaviour was maintained throughout the rest of the programme.

INTRODUCTION

The systematic use of learning principles in the treatment of behaviour disorders has resulted in the application of these principles to the child's natural environment. If most of a child's behaviour is maintained by its effects upon the natural environment, then it can be most effectively modified by changing the reinforcing contingencies supplied by the social agents in constant interaction with the child, e.g., parents, teachers. This approach assumes that modification of the child's natural environment is necessary for the maintenance of newly developed adaptive behaviour.

Encopresis, or fecal incontinence, is a lack of bowel control. Etiology is varied and may be related to toilet-training habits, parent and family orientation and experience, low intelligence of the child, age, poverty and large families, psychologically related conflict, stress, insecurity and over-protection, regression, and anal fixation, the expression of aggression or resentment, and attention-seeking behaviour (De Souza *et al.*, 1975).

In treating encopresis, if the component skills are within the child's behavioural repertoire but are displayed inconsistently, the objective is simply to increase the frequency of appropriate toileting behaviour. Neale (1963) and Gelber and Meyer (1965) did just this by immediately reinforcing appropriate toileting behaviour. Tomlinson (1970) also showed that by manipulating reinforcement contingencies it was possible to acquire control of voiding. Ayllon *et al.* (1975) have tried to make use of the verbal repertoire of the child in conjunction with the technique of reinforcer sampling. In the majority of cases, however, these studies were carried out in institutional settings.

The subject of this study was a severely deprived 8-year-old boy of limited intelligence who defaecated inappropriately. The child had always been totally encopretic, though not enuretic, despite the efforts of the professional services who helped the family. Medical examination revealed that there was no organic cause for the problem.

Two factors led to the decision to carry out the programme within the classroom, viz. :

(1) According to mother, the child was not encopretic at home though the validity of this statement was always held to be doubtful.

(2) The teacher's negative attitudes towards the child had increased to the point where she was becoming quite disturbed by the unpleasantness of the situation and frustrated in the time spent coping with the problem.

The clinical and educational psychologist worked together in the classroom with the teacher and the child in an attempt to solve what had become a major problem.

CASE HISTORY

David was brought to the attention of the Child Guidance Service when he was 7 years old. In the infant school there were several difficulties: his behaviour was disruptive, his ability to cope with the early learning activities was deficient and encopresis was a problem. Since assessment revealed that David's IQ was within 65-70 range, he was transferred to special school where his behaviour was brought under control. He was able to cope with the learning tasks, but his encopresis showed no signs of improving.

David's background was one of severe deprivation. Because of mother's limited intelligence and apparent inability to look after children, the Social Work Department assumed the responsibility of removing her children from her at birth. Most of her children had, therefore, been fostered in infancy though she had been allowed to keep two children, David and an infant of 12 months. The level of deprivation in the family was high though the Social Work Department was actively involved in supporting them.

David had been toilet trained for bladder activity at 3 years and there was no problem of enuresis. He was a deprived child who craved for attention and we believed that this was a causative factor in his encropresis.

METHOD

I. Weeks 1–3. The first phase of this programme involved establishing a baseline for weekly soiling and appropriate use of the toilet. The frequency of soiling in one week was high (Fig. 1). David arrived at school each morning having soiled himself on the bus en route to school. He soiled on average 10–12 times a week, usually after meals. The teacher responded to this by taking him to the toilet, cleaning him and changing his clothes.

II. Week 4–6. The time spent by the teacher in coping with the problem was re-organised so that her attention could be channelled appropriately. In our first attempt the teacher herself manipulated the reinforcement contingencies for appropriate behaviour. At intervals of one hour the teacher removed the child to the toilet. Cleanliness or correct evacuation was paired with social approval and praise. Soiling was paired with social disapproval and a show of anger. During this period no significant improvement occurred. Observation of the teacher's behaviour revealed the probable reasons for this failure. (1) Because of the unpleasant experience of having to handle the problem for so long previous to the setting up of this programme, she was unable to act in a discriminating way. In both the positive reinforcing and punishing situations voice modulation and attitude toward the child were noticeably similar. (2) In both situations David was receiving the same amount of attention and was unable to discriminate the positive from the punishing attitudes.

III. Weeks 7–15. In our second attempt at coping with the problem we introduced a full-time nursery nurse as the reinforcing agent. She was taught the use of differential reinforcement paired with a time-out contingency as mild punishment. This left the teacher free to respond to David in relation to his classroom activities only.

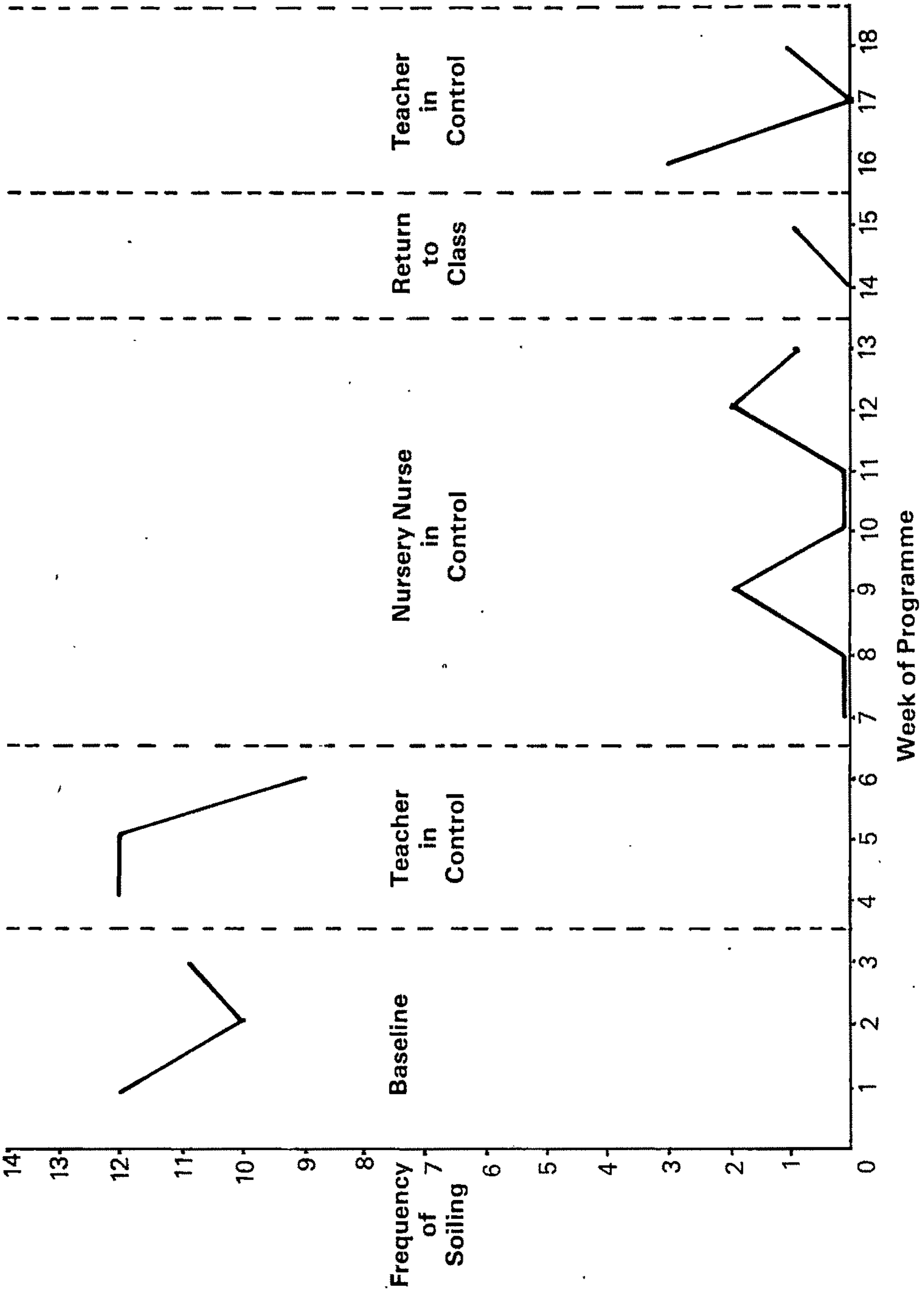
The programme was designed as follows :

(*a*) The nursery nurse met David at one-hourly intervals in the classroom beginning at 9 a.m., when he first arrived at school. He was taken to the toilet. Cleanliness or correct evacuation was again paired with social approval and praise, but he was also taken to a special room created for him. Here the nursery nurse reinforced David verbally for his cleanliness, played with him (painting, constructing models, etc.) and giving him her full-time attention. This session lasted approximately 30 minutes. The nursery nurse also encouraged David to stick stars on a prepared poster for clean behaviour.

(*b*) If David soiled himself between sessions the nursery nurse punished him by telling him he must clean himself. She then walked away, leaving him to care for himself and refusing to remove him for this reward session. This reaction by the nursery nurse made a deep impression on the child and his extreme encopretic behaviour ceased almost immediately.

FIGURE 1

Summary and Results of Behaviour Modification Programme



(*c*) The interval between sessions extended from one hour to after-meal times only. This proved more difficult for David (Week 12) though his level of responding remained much higher compared to the baseline.

(*d*) The nursery nurse gradually decreased the length of time she spent with David, first to 20 minutes, then to 10 minutes (Week 13).

(*e*) In the final stages (Weeks 14 and 15) the nursery nurse did not remove the child from the class but reinforced him within the classroom.

(*f*) Weeks 16–18. The nursery nurse was withdrawn from the classroom and the teacher, at this stage, was able to take over from her.

DISCUSSION

Our target had been to obtain an increase in the rate of self-initiated elimination behaviour to the operant level of the average child's eliminative behaviour. This was not achieved completely but it was possible to achieve a level of control which made the child's problem much more tolerable to the teacher and to himself.

The results of this case study indicate that the combined procedure of verbal instructions and reinforcer sampling was effective in rapidly increasing the appropriate toileting behaviour of the child. There appeared to be two characteristics of the child which seemed to be pre-requisite to the success of this type of treatment, viz.: (1) all the toileting components were within David's behavioural repertoire, and (2) he had sufficient verbal competence to circumvent the necessity of shaping behaviour and delivering reinforcement immediately after a response. The nursery nurse provided the instructions which served to link the target behaviour with the reinforcement contingencies. As Ayllon and Azrin (1964) have demonstrated, instructions, to be effective on a long-term basis, must be backed up by actual contingencies with which the subject comes into contact.

There seems to be little doubt that behavioural techniques can be effectively applied to children's problem behaviours within their natural environment. It involves the direct and practical coping with the everyday realities of the disorganisation resulting from the child's problem. In addition to therapeutic value, such an approach, which utilises the resources of the community, home or school, has important implications for future use in a systematic and preventive oriented model of mental health intervention.

ACKNOWLEDGMENT—The author wishes to thank Mrs. Ruth Symons, Senior Clinical Psychologist, Royal Aberdeen Children's Hospital, for her guidance and support throughout this programme.

REFERENCES

AYLLON, T., and AZRIN, N. H. (1964). Reinforcement and instructions with mental patients. *J. exp. Anal. Behav.*, 7, 327-331.

AYLLON, T., SIMON, S. J., and WILDMAN, R. W. (1975). Instructions and reinforcement in the elimination of encopresis—a case study. *J. behav. Ther. expt. Psychiat.*, 6, 235-238.

DE SOUZA, D. A., and DE SOUZA, A. (1975). Encopresis. *Child. Psychiat. Quart.*, 8, 1-8.

GELBER, H., and MEYER, V. (1965). Behaviour therapy and encopresis: the complexities involved in treatment. *Behav. Res. Ther.*, 2, 227-231.

NEALE, H. (1963). Behaviour therapy and encopresis in children. *Beh. Res. Ther.*, 1, 139-149.

TOMLINSON, J. R. (1970). The treatment of bowel retention by operant procedures: a case study. *J. beh. Ther. expt. Psychiat.*, 1, 83-85.

(*Manuscript received 22nd October*, 1976)

Br. J. educ. Psychol., 47, 203-208, 1977

CANDIDATE FATIGUE AND PERFORMANCE ON GCE OBJECTIVE TESTS

By A. J. MASSEY

(*Test Development and Research Unit, Oxford Delegacy of Local Examinations, University of Cambridge Local Examinations Syndicate, Oxford and Cambridge Schools Examination Board, Cambridge*)

SUMMARY. Four multiple choice tests were set to groups of between 80 and 103 Advanced level chemistry students in their original form or with the items in reversed order. Analysis of the pattern of responses showed that the students seemed able to complete these tests in the time available. Analysis of covariance, using level of achievement (as measured by score in an A-level multiple choice test) as the covariate, of scores on items in original and reversed order forms, showed no evidence of differences in the proportions of students choosing the correct response which might be attributed to the effects of candidate fatigue or time limitations.

INTRODUCTION

It has long been recognised that students' performance on test questions may be affected by structural aspects of the test itself. Willmott and Hall (1975) have shown how, in tests involving question choice, performance declines on questions completed late in the test.

Performance towards the end of the test may be inhibited by the onset of fatigue, by the effect of time limitations being most powerful late in the test, or by a combination of these. Whilst regarding these problems as significant in the context of non-objective tests Willmott and Hall did not investigate their effects in objective tests. They point out that most objective tests used in the General Certificate of Education (GCE) examinations are 'power' tests, completed by most candidates, rather than speeded tests where rate of working would have an effect on the scores obtained. Davidson and Carroll (1945) have shown that speed and power factors may be independent. Myers (1952) suggests that power may be more valid than speed as a predictor of academic performance, providing support for the GCE strategy of employing ' power ' tests. It would, however, be a mistake to ignore the potential effects of fatigue on items placed late in an objective test. If performance on such items was depressed it could have implications for the length of the tests which should be set and the ordering of items. Test constructors may not wish to measure differences in the stamina of students in subject examinations, or wish to avoid the situation where items testing particular topics are grouped at the end of a test, where fatigue effects may occur. The design of pre-testing programmes intended to ascertain the characteristics of items prior to their operational use would be affected, as would the interpretation of item analysis statistics.

In a study intended to illuminate the effects of speededness and item placement on item analysis statistics for College Entrance Examination Board (CEEB) tests in the USA, Mollenkopf (1950) has provided relevant evidence. Under conditions where every respondent was allowed sufficient time to complete the tests, he found some evidence of higher performance on verbal antonyms items when these were placed at the beginning of the test, rather than at the end. He found no such evidence on a test involving (more complex) mathematics items. Aiken (1964) investigating the effect on item performance of the difficulty of a 25-minute test which provided the item's context found item position immaterial.

The type of item usually set in a GCE examination is relatively complex by comparison with those used by Mollenkopf or Aiken and GCE tests are usually rather longer in terms of time allowance, although they may contain a similar or smaller number of items. Thus the possibility that fatigue may affect performance appears worth further investigation. If some candidates suffer from fatigue or have difficulty finishing, items placed late in a test would appear more difficult (exhibit a lower facility value—p) than if positioned early in the test.

This study investigates these possibilities by (*a*) observing the extent to which items positioned towards the end of GCE-style multiple choice tests are omitted (since omission is likely to be due to speed/time limitations), and (*b*) testing the hypothesis that there is no difference in performance on items when these are placed early or late within a test (i.e., the effect due to fatigue or to students being forced to guess by shortage of time).

PROCEDURES AND RESULTS

The study used multiple choice items included in the pre-test programme for the Advanced level examination in chemistry of the Oxford and Cambridge Schools Examination Board. Four 31-item pre-tests (Tests 1–4) were constructed, each including a relatively high proportion of items involving complex tasks. Two forms of each pre-test were printed. In the second form (prefixed R) the order of the items was reversed, item 31 becoming item 1, and so on. Schools were asked to administer the tests under simulated examination conditions. The time allowed was 1¼ hours, which was intended to be sufficient for students to finish. The number of schools and the number of students completing each test form are given in Table 1.

Scores obtained subsequently by these students on the 50-item multiple choice paper (GCE/MC), which formed a part of the Advanced-level GCE chemistry examination, provide a means of comparing the levels of achievement of the groups sitting the original and reversed forms.

Results.

The means and standard deviations of scores on each experimental test form, estimates of reliability given by Kuder and Richardson's Formula 20 (Guilford and Fruchter, 1973, p. 416), means and standard deviations of scores on the GCE/MC test obtained by the groups of candidates, and product-moment correlations between total scores on the experimental and the GCE/MC test, are reported in Table 1.

TABLE 1

SUMMARY STATISTICS AND INTERCORRELATIONS BETWEEN EXPERIMENTAL AND GCE/MC TESTS.

Test Form	1	R1	2	R2	3	R3	4	R4
Mean experimental test ..	13·42	12·67	15·52	14·52	11·77	11·03	14·00	13·76
S.D. experimental test ..	4·41	3·58	4·32	4·41	5·65	5·84	3·46	3·63
KR 20 experimental test..	0·69	0·51	0·70	0·71	0·82	0·83	0·51	0·58
Mean GCE/MC	35·90	33·68	36·88	36·24	38·31	35·93	35·42	36·87
S.D. GCE/MC	9·97	9·21	8·59	9·33	9·21	9·51	8·83	8·58
Product-moment correlation between experimental and GCE tests	0·60	0·66	0·57	0·67	0·57	0·31	0·56	0·72
(No. of schools)	(4)	(5)	(3)	(4)	(4)	(4)	(5)	(5)
Number of candidates....	101	103	88	80	88	96	87	90

Scrutiny of the proportion omitting the last 10 items in each form of the tests (not reported here), compared with the proportion failing to respond to the same item in the reversed version, suggests that students were able to attempt all items within the time allowed. The number of candidates failing to respond was extremely small, ranging between 0 per cent and 3 per cent (or 0–3 candidates) and no consistent pattern of omissions related to the positioning of the items in the tests was evident. It seems probable, therefore, that time constraints were not an important factor in these tests.

To examine the hypothesis that there are no differences in the level of performance caused by the position of the items in the test, it is necessary to examine differences in the proportion of candidates of a comparable ability level choosing the correct option to items placed early in one test form and late in a reversed version.

In order to make allowance for variations in level of achievement (as measured by GCE/MC score) the technique of analysis of covariance was employed. ANCOVA assumes a linear relationship between score on each item and GCE/MC score. Tests for curvilinearity (Guilford and Fruchter, 1973, p. 290) were applied (for each form) and in no instance was the proportion of variance accounted for by a curvilinear model significantly greater than that explained by a linear model. It was thus possible to pursue the ANCOVA. This analysis was carried out item by item using the multiple regression procedures of SPSS (Nie *et al*, 1975), and may be viewed as a method of asking the (sequential) questions 1, 2 and 3 below.

(1) Does knowledge of GCE/MC score together with knowledge of the version of the test help us to predict score on the item?
(Note that if the answer to question (1) is negative, the analysis cannot be continued.)

(2) Is the relationship between achievement in GCE and performance on the item the same whether the item is early or late in the test?
(Note that if there is an interaction effect, a lack of homogeneity of slopes, the analysis cannot be continued.)

(3) Do groups of students presented with the item late as opposed to early in a test, score differently after the effects of variations in the level of ability (measured by GCE/MC score) have been allowed for?

For each item regression procedures were used to obtain multiple correlations giving:

(i) The proportion of variance explained by using GCE/MC score as a predictor of score on the item (R^2_X).

(ii) The proportion of variance explained by using test form as a predictor (R^2_Z).

(iii) The proportion of variance explained by using test form and GCE/MC score as a predictor ($R^2_{X, Z}$).

(iv) The proportion of variance explained by using test form, GCE/MC and including the interaction term as a predictor ($R^2_{X, Z, XZ}$).

The ANCOVA table is given in Table 2, where, as an illustration, the analysis for item 1 in test 1 (31 in R1) is provided.

Thus the test for question 1 is a test of whether the saturated regression model has any predictive power. If it does not then subsequent steps in the analysis cannot be made and later variance ratios in the sequence have not been calculated. As we would expect in tests using previously untried items a substantial minority of items fail this test. Reference to item analyses using the point-biserial correlation between item and total test score (not reported here) confirmed that these items were failing to discriminate. Items which fall into this category are listed in Table 3 (where numbers of items are those of the original versions of the tests, i.e., 1, 2, 3, 4).

The test for question 2 examines interaction between the form of the test and GCE/MC scores. It asks whether the slopes of the regressions for the two test forms may be regarded as parallel. In the case of a small number of items (see Table 3) this was not the case, suggesting that the level of achievement of the candidate did not have the same effect on performance when the item was placed in different test forms. The positions occupied by these items does not suggest that this was related to the effects of fatigue. For these items the subsequent step in the analysis would be inappropriate.

This final step answers question 3 by asking if students (of a given level of ability) completing one test form have scored as well on the item as did those taking the reversed form. Thus it asks whether the regressions for the two test forms may be regarded as being separate from, rather than coincident with, one another.

TABLE 2

ANALYSIS OF COVARIANCE AND ILLUSTRATIVE EXAMPLE (ITEM 1 OF TEST 1).

Sources of variation	Sums of squares (values for Item 1 of Test 1)	df	Variance Ratio (values for Item 1 of test 1)	Question
(1) SS due to Z and X saturated model	$R^2_{x,\,z,\,xz}$ (= 0·06031)	3	$\frac{(1)/df(1)}{(4)/df(4)}$ (=4·299**)	1
(2) SS due to interaction (lack of homogeneity of slopes)	$R^2_{x,\,z,\,xz} - R^2_{x,\,z}$ (=0·06031—0·0567 = 0·0036)	1	$\frac{(2)/df(2)}{(4)/df(4)}$ (=0·770)	2
(3) SS due to Z adjusted for X	$R^2_{x,\,z} - R^2_{x}$ (0·0567—0·04988 = 0·0069)	1	$\frac{(3)/df(3)}{(4)/df(4)}$ (=1·454)	3
(4) SS residual	$1 - R^2_{x,\,z,\,xz}$ (=1—0·06031 = 0·9397)	N-3 (204-3)		

**=·01 significance level.

In the case of only 13 items from these four tests is the difference significant. If the hypothesis that placing an item late in the test will result in a lower facility for the item than if the same item were to be positioned early in a test is to be supported, we should look for significant differences in the items at each extreme of the order of presentation. We should expect ' early ' items in the original forms (tests 1–4) to have higher facilities than the same items in the reversed forms (tests R1—R4), whilst ' late ' items in the original forms (which came early on the reversed forms) would be expected to show higher p values for forms R1—4 than for forms 1–4. The 13 items where significant differences were observed and whether or not these differences were in the hypothesised direction are reported in Table 3.

TABLE 3

SUMMARY OF ANALYSIS OF COVARIANCE RESULTS.

Test forms	Items failing on Q.1	Items passing on Q.1 but failing on Q.2	Items passing on Qs. 1 and 2 but failing on Q.3	Items where difference between forms is significant after allowance for covariate
1 (N=101) and R1 (N=103)	2, 3, 7, 10, 16, 17, 18, 20, 25, 28	15, 23	1, 4, 5, 6, 8, 9, 11, 13, 14, 27, 29, 30	12*(A), 19**(X), 21*(X), 22*(X), 24*(A), 26*(X), 31*(A)
2 (N=88) and R2 (N=80)	4, 5, 6, 7, 10, 11, 15, 24 25, 26, 27, 30	8, 19	1, 2, 3, 9, 12, 13, 16, 17, 18, 20, 21, 22, 23, 28, 29, 31	14*(A)
2 (N=88) and R3 (N=96)	1, 6, 7, 10, 11, 17, 19, 20, 28, 29, 30	8, 9, 14, 21	2, 3, 4, 12, 15, 22, 23, 24, 25, 26, 27, 31	5*(A), 13*(X), 16*(X), 18*(X)
4 (N=87) and R4 (N=90)	10, 11, 12, 14, 16, 19, 20, 21, 25, 26, 27, 28, 30, 31		1, 2, 3, 4, 5, 6, 8, 9, 13, 15, 17, 18, 22, 23, 24, 29	7*(A)

*= ·05 level
**= ·01 level

(A)=difference in hypothesised direction
(X)=difference in opposite direction to that hypothesised.

No early items (1–10) in form 1 showed significant differences. Of the late (22–31) items four showed significant differences between forms. Items 24 and 31 did show higher facility values in the reversed version but the differences for 26 and 22 are in the opposite direction (higher facility values when positioned towards the end). In test 2 no early or late items showed significant differences whilst in test 3 difference was apparent only for item 5, where the facility value was 0·502 as compared with 0·427 as item 27 in test R3. In test 4 item 7 had a facility value of 0·598 as compared with a value of 0·467 as item 25 in R4 and was the only item showing a significant difference.

DISCUSSION

There is no evidence that the students sitting these 1¼ hour tests showed any decline in performance at the end of the tests. This would support previous research suggesting that performance on items embedded within tests composed of more complex items and administered with an adequate time allowance, thus testing power as opposed to speed, is unaffected by the positioning of the items within the test. The tests employed in this study were similar to those in common use in British examinations.

The time allowance for the tests was similar to those usually encountered and there would, therefore, appear to be no evidence supporting concern that fatigue is a factor in the performance of candidates on multiple choice tests (of this length) forming part of GCE examinations, or for fears that fatigue or time constraints are affecting responses in pretests. This would not, however, preclude the possibility that fatigue may affect performance in the relatively small number of objective tests with time allowances in excess of 1¼ hours in use in British examinations. Further studies would be necessary to establish the role of fatigue in such situations.

REFERENCES

AIKEN, L. R. (1964). Item context and position effects on multiple choice tests, *J. Psychol.*, 58, 369-373.

DAVIDSON, W. M., and CARROLL, J. B. (1945). Speed and level components in time scores: a factor analysis. *Educ. psychol. Measur.*, 5, 411-427.

GUILFORD, J. P., and FRUCHTER, B. (1973). *Fundamental Statistics in Psychology and Education* (5th edit.). New York: McGraw-Hill.

MOLLENKOPF, W. G. (1950). An experimental study of the effects on item analysis data of changing item placement and test time limit. *Psychometrika*, 15, 291-315.

MYERS, C. T. (1952). The factorial composition and validity of differently speeded tests. *Psychometrika*, 17, 347-352.

NIE, N. H., HADLAI HULL, C., JENKINS, J. G., STEINBRENNER, K., and BENT, D. H. (1975). *SPSS—Statistical Package for the Social Sciences* (2nd edit.). New York: McGraw-Hill.

WILLMOTT, A. S., and HALL, C. G. W. (1975). *O-level Examined; the Effects of Question Choice.* London: Schools Council Research Studies, Macmillan Education.

(*Manuscript received 29th September,* 1976)

Br. J. educ. Psychol., 47, 209-214, 1977

THE ROLE OF INTERVIEWS IN CHOICE OF UNIVERSITY BY PSYCHOLOGY UNDERGRADUATES

By C. V. NEWMAN, R. COCHRANE and D. E. BLACKMAN
(*Department of Psychology, University of Birmingham*)

SUMMARY. Undergraduate applicants for a university degree course in psychology received conditional offers of places having undergone one of four different introductions to the department ranging from a group open day to a stressful individual interview with two admissions tutors. The prediction that as interviews became more stressful so the acceptance rate of conditional offers would increase was not confirmed. Instead, evidence suggested that an applicant's initial preference ranking of the five universities to which he had applied remained the salient factor influencing his subsequent decisions about accepting offers of places.

INTRODUCTION

A number of recent publications, such as that by Weir (1976), have been concerned with the ability of selection interviews to predict future academic success among applicants for undergraduate degree courses at university. Interviews of university applicants can, however, serve a different purpose, that of informing applicants about the university and encouraging them firmly to accept any places offered them. Indeed, under the present system of entry to British universities conducted by the Universities Central Council for Admissions (UCCA) interviews are often conducted with this goal rather than genuine selection in mind. Applicants for a university place are permitted to apply simultaneously to five universities, listing them in order of choice, though equal preference rankings are allowed. Usually around February, universities make offers conditional on the candidate gaining certain specified grades in their forthcoming GCE A-level examinations, held around July. Where an applicant is offered a conditional place at more than one of the five universities to which he has applied, he is required to make a choice, firmly accepting one offer, provisionally accepting a second (usually less demanding) offer and declining all other offers.

Candidates' decisions at this stage in the admissions procedure are not bound by their initial preference ordering of the universities to which they have applied. It is quite possible for candidates eventually to accept firmly a conditional offer from the university they ranked fifth on their original application and to decline a similar offer from the university they initially ranked as their first choice. In practice, therefore, although universities are selecting candidates, the candidates are in turn selecting universities from amongst those that make them conditional offers of places. Many factors may influence a candidate's eventual choice, but his experience at interview may be one of the most important factors determining his attitude towards a particular university and hence his eventual decision to accept or reject a conditional offer of a place.

The present experiment was designed to investigate whether different types of interview or similar introduction to a department had any influence on applicants' eventual acceptance or rejection of a conditional offer made by the university department. The chief aim of the study was practical: to provide factual data on which to base decisions concerning admissions procedures. However, a guiding hypothesis was provided by cognitive dissonance theory (Festinger, 1957). From this theory it would be predicted that a candidate who is made to feel that he has earned his conditional offer of a university place only after a stressful and demanding interview experience will be more likely to value, and hence accept, this offer than will a candidate who perceives the receipt of a conditional offer to have been entirely without stress and virtually automatic. The four interview conditions described below were intended to give a graded series of increasingly stressful and demanding situations in which to test this hypothesis experimentally.

METHOD

Subjects for the study were selected from all candidates who applied for the single subject, B.Sc.(Hons.) degree in Psychology at the University of Birmingham during the years 1974 and 1975. Included in the study were all candidates who met the criteria of being final year sixth form students about to take three A-levels for the first time and having sufficiently good confidential references and academic records for the admissions tutor to feel able to make them a standard conditional offer of three C grade A-levels. These applicants were assigned in rotation, in order of receipt of their application forms, to one of the four interview conditions described below. As a result of certain candidates withdrawing completely from the UCCA application procedure at a later date, and the inability of certain candidates to attend for interview as a result of a national railway strike, a number of subjects had to be dropped from the study leaving 194 subjects, divided somewhat unevenly between the following four groups:

Group 1, *Open Day.*

Candidates in this condition were sent a letter informing them that through UCCA the University was about to make them an offer of a place conditional on their obtaining grades CCC in their three A-level subjects. In order to help them make an informed decision on this offer, the Department of Psychology was inviting them to attend an open day on a named date a few weeks ahead. As candidates were under no compulsion to attend this open day and as they knew they had already secured a conditional offer of a place at Birmingham University, this condition was deemed to be a non-demanding and unstressful admissions condition.

Numbers of candidates attending particular open days varied but usually about 40 applicants (the majority being candidates not included in the study) met in a lecture room where they were addressed informally for about half-an-hour by the admissions tutor before touring the department and university campus under the guidance of undergraduate volunteers.

Group 2, *Questionnaire and Open Day.*

Candidates assigned to this condition were initially sent a letter, the opening paragraph of which read as follows :

" The Department of Psychology is carefully considering your application for admission. Prior to making a final decision we would like to ask you to complete and return the enclosed form. Your answers to the questions on this form can in no way adversely affect your chances of receiving an offer of a place at this University. The form may, however, be used should the situation arise in which you receive an offer on the basis of A-level grades which you subsequently fail to achieve. In this case we may take into account your answers in deciding whether to make a further offer." The letter was accompanied by a questionnaire on which candidates were asked to answer six, rather general, non-searching questions such as " How did you first become interested in psychology? " and " Do you have any specific career ambition? " The candidate was finally asked to write briefly on what he expected to gain from studying psychology.

After two weeks, and whether or not the questionnaire had been returned, all candidates were sent the same letter as those in Group 1 inviting them to an open day and informing them that a conditional offer of CCC was being made them. Open days for Group 2 candidates were identical to those in Group 1.

Ethical considerations prevented us from allowing the candidates to believe that the answers to the questionnaire formed an integral part of our selection process; however, the fact that a candidate was asked to spend time filling out a form necessitated some degree of effort on the part of a candidate before a conditional offer of a place was received. Despite assurances to the contrary, the delay between receipt and return of the form and receipt of the eventual offer may have generated some slight feeling of uncertainty that the latter event was in some way contingent on the manner in which the questionnaire was completed. Certainly candidates' responses indicated that care had been taken over the answers.

Group 3, *Single Interviewer.*

In this group candidates were sent a letter the opening paragraph of which was identical to that sent to the Group 2 candidates except that they were asked to attend for an interview rather than fill in a form. Those candidates who accepted the invitation were interviewed by one of the three authors of this report who each saw a third of them and, at interview, tried to create a relaxed, friendly and informal setting in which all candidates were asked the same six non-searching questions as had been included on the questionnaire sent to Group 2 candidates. Some factual information on the department was given to each candidate and any queries answered. The whole interview lasted about 15 minutes after which candidates were invited to tour the department and the campus. All candidates received the standard conditional offer a few weeks later via UCCA.

In this condition the interview was deliberately made informal and without stress. The candidate did, however, have to invest a certain amount of time and effort in travelling to Birmingham and appearing at interview. Despite an assurance to the contrary in the letter of invitation to the interview, it is possible that candidates believed that their performance at interview in some way influenced their eventual receipt of an offer. Nothing was done deliberately to dispel any such beliefs.

Group 4, *Two Interviewers.*

Candidates assigned to this condition were sent the identical letter of invitation to those in Group 3; however, on arrival at the department they were interviewed simultaneously by two of the authors of this report who asked questions in turn. Each of the three interviewers was involved in two-thirds of the interviews. A more demanding atmosphere was created in which both interviewers remained serious and formal. Initially the candidate was asked the same six unsearching questions as those in Group 3, but these were followed by five pre-determined and more difficult questions requiring some knowledge of psychology to answer: e.g., " What does the term ' conditioning ' refer to?", " What do you know about the work of Freud and Eysenck?" Brief and evasive answers were pursued by deliberately probing questions so that even the more self-assured or well-informed candidates were faced with the inability to answer certain questions. The interview was terminated as for Group 3 candidates and conditional offers were again made to all candidates within a few weeks of the interview via UCCA.

Although all other conditions were identical to those of Group 3 candidates, the interview itself was considerably more stressful. Most candidates showed signs of agitation and embarrassment on being unable to answer two or three successive questions and none of the interviewers was left in any doubt that they had created a stressful and somewhat demanding interview situation in which candidates appeared to be genuinely 'on test.' In accordance with the experimental hypothesis, this condition should have evoked the greatest number of firm acceptances of conditional offers made to the candidates.

RESULTS AND DISCUSSION

By June of the year in which candidates hope to enter university in the following October, they must have responded in one of three ways to the conditional offers made to them by the different universities to which they have applied: firmly accept (FA); provisionally accept (PA); or decline (D). Table 1 shows how candidates in the four interview conditions responded to the conditional offer of a place to read for a degree in psychology at the University of Birmingham.

The effect of interview conditions on candidates' decisions can be seen to be slight and χ^2 tests revealed there to be no statistically significant differences in the distribution of responses between the four conditions. There is, therefore, no evidence to confirm the experimental hypothesis that a demanding interview increases the number of candidates accepting conditional offers made to them.

TABLE 1

NUMBER OF CANDIDATES RESPONDING TO CONDITIONAL OFFERS AND ATTENDANCE FIGURES AT INTERVIEW OR OPEN DAY.

Group	Categories of Response	Number of Candidates per Response	Number of Candidates as % of total in group	Number of Candidates who attended the interview or open day
Group 1, Open Day (N=53)	F.A. P.A. D.	20 6 27	37% 12% 51%	14 2 11
	Total	53		27
Group 2, Open Day and Questionnaire (N=54)	F.A. P.A. D.	26 10 18	48% 19% 33%	19 6 8
	Total	54		33*
Group 3, One Interviewer (N=43)	F.A. P.A. D.	17 9 17	39% 22% 39%	16 9 13
	Total	43		38
Group 4, Two Interviewers (N=44)	F.A. P.A. D.	19 9 16	43% 21% 36%	19 9 12
	Total	44		40

* 44 of these candidates returned the completed questionnaire.

It was also possible to reject a subsidiary hypothesis that a visit to the university increases the likelihood that a candidate will accept a conditional offer of a place. Although 89 per cent of all candidates invited for interview actually attended as against 56 per cent of those invited to an open day, 41 per cent of all candidates in the two interview conditions firmly accepted conditional offers compared with 42 per cent of the candidates in the combined open day conditions. This suggests that a forced visit to the university to attend an interview does not generate increased interest in the university. Of course, this result does not mean that open days and interviews serve no purpose. From Table 1 it can be seen that in all groups a considerably higher proportion of candidates who eventually firmly accepted offers actually visited the university than did candidates who declined offers. This outcome would be expected but, without a fifth experimental condition in which candidates were offered places with no invitation to visit the university, the importance of a visit itself cannot be established conclusively.

Further examination of the data provided evidence that a candidate's initial preference for a particular university was the salient factor in determining his or her final choice so that other factors such as interviews and visits to universities have only a minor influence on eventual decisions. Amongst all subjects irrespective of experimental conditions 61 per cent of those who had nominated Birmingham as their first or second choice firmly accepted the offer they received. Of those candidates whose initial choice put Birmingham lower than second only 26 per cent firmly accepted the offer. The same conclusion follows from the data shown in Table 2 where the acceptance rate of offers is shown for applicants in the different experimental conditions.

Though not statistically significant, the data in Table 2 suggest the possibility that the open day condition was less effective than all other conditions in evoking firm acceptances from candidates with initially a low preference for the university but was as effective as all other conditions in attracting firm acceptances from candidates initially ranking the university highly. This possible interaction between interview experience and initial preference should be investigated further.

TABLE 2

NUMBER OF CANDIDATES AND PERCENTAGE OF TOTAL (IN PARENTHESES) WHO RESPONDED TO CONDITIONAL OFFERS IN TERMS OF INITIAL PREFERENCE ORDER AND EXPERIMENTAL GROUP.

Group	Initial Preference Order	Total	F.A. (%)	P.A. (%)	D. (%)
Group 1	1 or 2	28	17 (62)	3 (10)	8 (28)
Open Day	3, 4 or 5	25	3 (12)	3 (12)	19 (76)
Group 2 Open Day+	1 or 2	26	16 (62)	4 (15)	6 (23)
Questionnaire	3, 4, or 5......	28	10 (35)	6 (23)	12 (42)
Group 3	1 or 2	20	11 (55)	2 (10)	7 (35)
One Interviewer ..	3, 4 or 5	23	6 (26)	7 (31)	10 (43)
Group 4	1 or 2	16	11 (68)	2 (13)	3 (19)
Two Interviewers ..	3, 4 or 5	28	8 (29)	7 (25)	13 (46)

CONCLUSIONS

Candidates selecting psychology as a future degree subject seem to have decided on their preferred universities before they fill out their application forms. Having decided their initial preference order between the five universities of their choice, later experiences at interviews have relatively little effect as far as reappraisal of their initial preferences for different universities are concerned. Once a candidate has declined a conditional offer from a university, UCCA does not send that university any further information on the candidate. It was, therefore, not possible to investigate the influence of other variables such as the level and number of other offers made to candidates on their eventual decision to accept or decline the conditional offer they received from the University of Birmingham. It is probable that, in practice, this factor is not very important as the UCCA procedure that allows candidates to hold an offer provisionally means that a candidate can firmly accept a stiff offer from the university he really prefers most and provisionally accept a low, 'safety net' offer from a different university. It was assumed that any influences from such variables as the above were likely to act at random between the four experimental conditions of the study.

The inconclusive finding that type of interview had no significant effect on candidates' choices may have arisen because the perceived significance of the experimental conditions had been reduced by the assurance given to all candidates concerning the lack of emphasis being placed on the interview or questionnaire in the total admissions procedure. Research conducted in real life situations inevitably places certain ethical restraints on the manner in which such variables as those considered in the present study can be manipulated. It could be argued that although subjects in the different experimental groups underwent different degrees of stress at interview, the effect of this variable had been modified by the dissociation created between interview conditions and eventual outcome of the candidates' applications. This interpretation of the results cannot be dismissed, but on the other hand, a significantly higher percentage of interview candidates came to Birmingham than those invited for open days (Table 1). Presumably this occurred because candidates did assume we were placing some real importance on the interview despite the single assurance to the contrary. In practical terms, however, the results might be taken to suggest that the time and effort spent in a department of psychology interviewing applicants individually for other than genuine selection purposes could be more economically devoted to organising group interviews or open days for applicants or potential applicants who have yet to apply to university.

REFERENCES

FESTINGER, L. (1957). *A Theory of Cognitive Dissonance.* New York: Harper and Row.

WEIR, C. (1976). Interviews as an aid to selection of psychology undergraduates. *Br. J. educ. Psychol.*, 46, 88-93.

(*Manuscript received* 24*th November*, 1976)

BOOK REVIEWS

CLARK, M. M. (1976). *Young Fluent Readers.* London: Heinemann Educational, pp. 116, £1·70.

For some years, Dr. Margaret Clark, Reader in Educational Psychology in the University of Strathclyde, has attracted attention through her investigations into the reasons why some children fail in reading whereas others succeed. In *Young Fluent Readers* she looks closely at a group of 32 children who on starting school were already able to read at a level which placed them outside her category of potential educational failure. Her criterion for ' fluency ' in this instance was a reading age of 7½ years on the Schonell Graded Word Reading Test. Children were observed over a two-year period with particular reference to attainment, social class background and pre-school experience, the general aim being to discover how and why they had learned to read so well.

The sample was gathered with the help of a large number of schools but no attempt to match subjects with non-readers of similar intelligence and socio-economic background is reported. Nevertheless, it was possible to refer performance of the 32 children to that of gifted children from earlier studies. In the event, 21 of the children had IQs in excess of 140 (Stanford Binet) and most scored at a level two years ahead of actual age. Scores were also above average on the WPPSI with some children reaching item ceilings on a number of subtests and only four of the pupils having Performance IQs superior to Verbal IQ. In addition, excellent all-round attainment was reported with very few of the sample outstanding only in reading.

Dr. Clark maintains that success cannot be accounted for merely through reference to intelligence—there were other children of like ability who entered school unable to read. Nor is it the case, she insists, that pre-school experience alone is sufficient to explain progress. The more likely explanation is that the demands to seek out reading by a group of very able and determined children were met in such a way that learning was rapid and successful. Undoubtedly this is correct. But the study as reported is in no position to settle the issue one way or the other for without proper controls it is extremely speculative to specify aspects of ability which were responsible for the remarkable reading progress, or to spotlight those experiences which set the fluent readers apart. If the reader accepts this situation, and, in all fairness, Dr. Clark frankly admits that it was beyond resources to establish control groups, then there is much to fascinate in the study.

Teachers will be interested in the insightful comments on the usefulness of those tests given to children to help identify factors which influence learning to read, comments which clearly reveal a knowledge about children which many teachers feel psychologists lack. Psychologists will be interested in steps taken to improve the validity of tests which are applied to British children but developed in the USA, comments which reveal an extensive experience of testing which some psychologists feel teachers lack. Teachers and psychologists ought to heed the advice about difficulties in designing remediation programmes based on profiles derived from batteries of different tests.

Two emphases in the study are most welcome and their promise of constructive help may appeal to the teacher; the focus is on success as opposed to failure in reading, and on the processes by which children actually learn to read as opposed to how they are taught. Unfortunately, such promise is largely unfulfilled since certain key issues remain insufficiently resolved. Those factors which account for children starting school already able to read are not adequately established and identified, nor are those which promote readiness for reading and increase the chances of early success in pupils' learning. At the same time, it is difficult to reconcile the research strategy used in the study; that is, waiting for subjects to display success before investigating the processes which have been involved. Perhaps this point is unfair in that it is difficult to suggest a more convenient tactic given the realities of school-based enquiry.

The reader is presented with some interesting discussion and reports of aspects of home and school experience which may influence learning to read, in particular, the reviewer would commend the section on interviews with parents. Dr. Clark stresses the need for caution in interpreting correlations, pointing out that it is not valid to reason that this or that factor has caused reading success (it might have been pointed out that it would be equally invalid to reason that reading has caused the factors!).

The book is well worth reading but the findings of the investigation far too inconclusive for this reviewer.

TERRY DOLAN.

JONES-DAVIES, C., and CAVE, R. G. (Eds.) (1976). *The Disruptive Pupil in the Secondary School.* London: Ward Lock Educational, pp. 122, £1·95.

" These are the times that try men's souls."

Tom Paine.

At last it is now fashionable to mention the disruptive adolescent in the contemporary scene. Until recently such references to violent behaviour in secondary schools would have been brushed aside with some glib cliché about " a minority of cases " or " the author tends to over-state the case." In the Crombie era, school unrest has arrived as a topic worthy of serious discussion.

This short book of 122 pages is at least the opening shot in a debate that will no doubt reverberate throughout the next decade. The editors, sensibly enough, are well aware of the limitations of their subject matter. Ronald Cave writes: " No foolish claim is made that this collection of papers will provide easy answers to the problems caused by disruptive secondary-school students " (p. 6).

Logically enough, few positive solutions emerge. Even so, it is with a feeling of disappointment that one reaches the final page without any really significant thoughts in mind. There are some highlights, of course, as in most collected papers; the contribution by Brian Davies is witty and incisive and is suitably stocked with laudable references. From a sociological perspective some apt points are made. But why not admit the truth of the situation in the first place, that we *do not know* how to cope with extreme school indiscipline any more than the Cabinet can cope with the miners ?

Apart from the editors there are nine contributors who, in the main, are practitioners, the possible exception being Professor Ray Jackson of the University of Malawi. His paper has the usual material that one has come to expect from the ' Culture of the School ' debate, Hargreaves, Lacey, Waller and Douglas. These studies contribute little to the study of the violent pupil and neither does Becker or Lauwerys. The work of Professor Ladd of Atlanta University, Georgia, would have been more to the point. On the other hand, those at the sharp end, such as Bernard Baxter in his ' Tackling the Disruptive Pupil,' seem to lack a certain breadth of vision and sociological awareness in their work. What is basically required in this field are some tangible research studies on which to mount the argument and to suggest future strategies. At the moment one seems caught between sociological generalities on the one hand, and school anecdotes on the other. This, amongst other reasons, is probably due to the fact that the more enterprising teachers leave the classroom whilst the more radical educationists have rarely been in it.

Yet this is a brave (although somewhat naïve) book and it would be unfair to criticise the various authors for not solving a problem that, in the present context, might well be incapable of solution. At the moment, we appear to be trapped between educational elastoplast and social engineering. After all, who wants to hear about Clausewitz in ' No-Man's-Land '?

What is required to analyse the problem of the disruptive secondary school pupil is, firstly, an analysis of the parameters of the school culture, secondly, a battery of micro-studies on classroom behaviour (some of an interactionist nature), thirdly, a cool, hard look at teacher behaviour and school organisation, and fourthly, a new paradigm that would embrace the school as a learning system in a community gestalt. At the moment there are too many people singing in the lifeboat without much idea of their personal salvation.

Finally, this book is not unduly meticulous about presenting sources. For example, what happened to Stott (1958) in B. J. Flaherty's contribution? Also, an index and full bibliography would have been useful. However, a good skirmish in a war of attrition. Where do we go from here?

RICHARD FARLEY.

KLAHR, D., and WALLACE, J. G. (1976). *Cognitive Development: an Information-Processing View.* Hillsdale, New Jersey: Lawrence Erlbaum Associates, pp. xii+244, £9·40, $15·95.

In this work Klahr and Wallace concern themselves with the development of human cognition. In particular, they seek a precise set of algorithms, each algorithm just sufficient to model successful human performances on one of an important sequence of well-defined Piagetian tasks, with a view to deriving from the differences between these algorithms the nature of the developing system. The algorithms are empirically tested against human performance by writing them as computer programs; the most primitive of which is calibrated against its human counterpart to provide a measurement instrument for the rest of the series. In the words of the authors: " We believe that the precise formulation of the various stages of cognitive development makes it much more likely that a similarly precise model of the transition mechanism itself will soon be obtainable," although in this book that mechanism is only outlined.

After an introductory discussion on the central psychological problems to be faced, the authors describe their chosen programming tool. The choice of programming system for any particular job is always difficult, and their choice of a productive system is a useful one not only for the ease with which an information-processing model of cognition may be mapped onto it, but also the ease with which that mapping may be understood by a programming novice. Klahr and Wallace give a useful primer for the programming language, and complement this with a very detailed account of their most primitive algorithm, which handles quantitative comparisons. With this, as with later production sets, the algorithm is given in full, and its action described in detail on representative data. The detail is such that an interested reader could, if he chose, work through the processes with his own input.

The authors then describe algorithms for quantification, class inclusion, conservation of quantity and transitivity of quantity, showing the deficiencies of each for the subsequent task and describing the degree to which each matches empirical measures on humans, both in terms of success in task, and rate of working. In transit to the most complex task, the cognitive model itself is extended (as well as the production algorithms acting upon it), leading finally to a demonstration of the requirement for, and the nature of, processing in the sensory buffers.

The changes in algorithms necessary for tasks of increasing complexity raise questions as to the nature of the processes which change the algorithms themselves. The authors address themselves to this topic in the final two chapters, describing there a less specific (but more unified) model in which these equilibrations could take place. At the same time, these chapters represent a challenging programme of work, and the authors highlight problems in many areas which need to be met before empirical testing on such a developmental model could even start.

The precision with which the authors describe their theory is a brave and laudable step forward, useful in its own right independent of the worth of the theory Useful, because this precision avoids the aggravating situation wherein an ambiguous theory ' grows to fit ' its users' intentions. Brave, because the attention it so rightly deserves, together with its testability, will ensure a turbulent future for it; as one presumes its authors hoped. The clarity of exposition of the book will aid this process, as it is self-contained enough to appeal to psychologists, artificial intelligence specialists and those interested in automatic studies.

PHIL ODOR.

LIPSITT, L. P. (Ed.) (1976). *Developmental Psychobiology: The Significance of Infancy.* New Jersey: Erlbaum; distributed by John Wiley, pp x+143, £7·65, $12·65.

This slim volume contains chapters by five different authors, each of whom also comments upon one of the other contributions. There is also a concluding overview by Jerome Kagan. The papers were originally presented at a symposium organised by the American Association for the Advancement of Science.

What is developmental psychobiology? The book's editor, Lewis Lipsitt, has made his reputation by conducting immaculately clean experimental studies which effectively constrain the human infant within the role of providing neat and tidy kymographic records of physiological or easily recordable physical responses. These are reported in the traditional style of conventional scientific data presentation, often using graphs with time or trials along the abscissa. The book starts on this tack with a paper by Campos claiming that heart rate recordings represent a sensitive and appropriate tool for the study of emotional development in the infant. Clearly, at this level infant psychobiology suggests a commitment to precision recording techniques, focused upon hard data measurements, as the approach most likely to lead us towards a deeper understanding of ' the significance of infancy '—this being the somewhat grandiose sub-title to which this book lays claim.

What else distinguishes this particular approach ? The infant psychobiologists also seem to be constrained to come down fairly firmly on the ' nature ' side of that old false argument which attempts to apportion weights to the relative contributions of nature and nurture in human development. The article by Daniel Freedman thus suggests firstly that Chinese American infants are more constitutionally placid than their Caucasian counterparts and then goes on to argue that characteristic sex differences are to be found in children's drawings despite obvious cultural differences in artistic style. Still on the theme of heredity-*versus*-environment, Sandra Scarr-Salapatek attempts to make the argument more up-to-date by reference to Waddington's concepts of canalisation and creeds. Disappointingly, however, the only evidence she presents for discussion relates to infant mental scale correlations comparing identical and fraternal twins. Given the poor predictive validity of baby tests, it is hardly surprising that this discussion fails to lead to any new conclusions.

A further distinctive feature of the psychobiological position is apparently a concern with reflexive behaviour in the infant and here Philip Zelazo puts forward the view that the continual exercise of certain primitive reflexive responses can lead to the precocious emergence of such achievements as sitting, standing and walking. In discussing the smiling response, Zelazo even suggests that it may be useful to think in terms of a conditionable ' cognitive reflex '! In support of this position he relies heavily on the work of Charles Super who studied the Kipsigis in Kenya and on Melvin Konner's reports on the Desert San, a primitive hunter gathering people living in the Kalahari Desert. Apparently both cultures have a strong belief in the need to be given practice in certain early motor skills and their infants are distinguished by reaching their motor milestones correspondingly early. Following this paper there is an interesting disagreement between Zelazo and Freedman over interpretation, but at the end of it all it still remains an open question whether the reported precocity of certain African babies is real or artifactual and the only conclusion finally arrived at is that more serious work on infant reflexology is urgently called for.

In his paper entitled ' Developmental Psychobiology Comes of Age,' Lipsitt first claims his editorial privilege to comment rather generally on what has gone before and he then attempts to discuss certain practical implications of his own work on neonatal sucking. His discussion ends with a section on the relevance of developmental psychobiology to crib death in which he concludes: " It is not inconceivable that a modicum of early practice is required for the infant to achieve a suitable level of respiratory retrieval when faced with occlusion." Perhaps understandably he fails to take this line of thought to its logical conclusion by advocating that infants be given practice in fighting for breath by allowing them the experience of being partially suffocated under carefully controlled conditions.

What then does developmental psychobiology add up to? Mainly, perhaps, an assertion that the conventional experimental approach with a special insistence on careful and precise techniques of measurement firmly anchored in the S–R tradition is still the only safe way forward. Lip-service is paid to the idea that development is also a two way interaction process and that ' context ' is somehow an important factor that needs to be taken into account, but that is all. For this reviewer the theoretical position of this book is distinguished mainly by an implicit assumption that, for the purposes of serious study, it is essential to treat a developing human organism exclusively as an experimentally manipulable' thing object' and never as a person with whom direct communication might be even a remote possibility.

JOHN NEWSON.

MALONEY, M. P., and WARD, M. P. (1976). *Psychological Assessment: A Conceptual Approach.* New York: Oxford U.P., pp. xiv+422, £11·50.

In this large 400-page book, Maloney and Ward, who write from the perspective of clinical psychologists, put the case for using all the various types of psychological tests and other data in psychological assessment. In effect, therefore, this book is a survey of current psychological tests and their rationale together with other forms of assessment. However, compared with many books on testing, it stresses the importance of using test scores as one, may be very important, source of data together with information obtained from interviews, life and medical records. Such a viewpoint, of course, neatly makes nonsense of many objective academic studies of projective tests where attempts are made, almost always without success, to score protocols blind.

From this brief description it is clear that this is a book mainly for those who are to use tests professionally (rather than the undergraduate). How well do the authors achieve their aims? After all there are many books about testing. For what reasons might we be persuaded to use this one?

The first point to notice about this book is that psychometrically it is not advanced or up-to-date. Thus the descriptions of reliability make no mention of any unusual varieties, such as analysis of variance methods. Similarly, their discussion of validity is elementary although they utilise a concept of conceptual validity in terms of assessing an individual. Rather more importantly there is no discussion of item-scaling procedures such as Rasch scaling. However, even if elementary the material is clear and well-written.

Their notion of conceptual validity becomes highly important when they begin their chapter on the assessment of personality. Indeed, that this book is not just another book on tests becomes clear. They write " The basic premise of this text is that psychological tests by themselves are inefficient tools for answering assessment questions and that such questions must be answered by a way of a conceptual model based on a hypothetical deductive reasoning approach " (p. 159). This model, the Personal Construct model, aims to predict individual behaviour based upon, as we would expect, data from all sources. A case study is then included to illustrate the method. Do the authors make their case that psychological testing alone is insffiucient to assess an individual? As the arguments are presented in this book, they do. However, the multivariate approach, as advocated by Cattell and Eysenck, is never adequately discussed and although, at present, this is not perfected, ultimately, perhaps, as the specification equation and the dynamic calculus demonstrate, it might be. Thus, their Personal Construct model is not wholly convincing.

Naturally, they devote a chapter to the concept of intelligence. This includes summaries of the learning theory position as put forward by Staats, the biological argument of Hebb, the psychometric position as exemplified by Guilford, together with Piaget's developmental approach. Although mention is made of Cattell's fluid and crystallised abilities, there is no discussion of Cattell's ADAC model which again makes this reviewer suspect that these clinical psychologists are not, as they say, into modern psychometrics.

This deficiency permeates their later discussions of personality tests. Although factor-analytic tests are mentioned the fundamental distinction between factor analytic variables and criterion keyed variables in terms of psychological meaning is never dealt with.

This then is the weakest side of the book. On the other side, however, the full discussion of cases, the sensitive use of test data together with case records and other information, the chapters on psychiatric diagnosis and brain damage make this book highly useful for the practical psychologist.

In summary we would conclude that this book is welcome in that it stresses the importance of being able to assess the individual, rather than discriminate groups and it will be helpful in professional psychology. As an academic psychometric text, it cannot be recommended.

PAUL KLINE.

SLATER, P. (Ed.) (1976). *Explorations of Intrapersonal Space, Volume I: The Measurement of Intrapersonal Space by Grid Technique.* London: Wiley, pp. 258, £8·75, $16·65.

One's first reaction to this book is one of curiosity that Volume 1 contains applications of a methodology which is, it seems, to be described in Volume 2. But at least this book gives examples of a wide range of applications for grid technique as a device for discovering what is going on behind the eyes that see, for example, a child's world, prison bars, seaside resorts, or events that turn the hand towards self-destruction. Some discussion is also spent in consideration of the geography of the environment proper, of the self and the correspondence between the person's and the therapist's view of each other.

Whilst the technique (which turns out to be several different ones, despite the Introduction) is originally derived from Kelly's personal construct theory, this reviewer at least was left wondering what the theoretical motor (and indeed theoretical outcome) is of some of the work reported. The chapters diverge and are diverse in their adherence to Kelly's theoretical beliefs and phrases like ' castration anxiety ' creep into the narrative at times. Some of the better chapters (Salmon's on children's thinking and Landfield's on suicide, for example) derive new and interesting conceptualisations of their respective concerns from Kelly's theory. The authors then turn, quite naturally, to the associated techniques to explore these insights. Other chapters employ Kelly-like views of thought processes but use them in informal and unstructured ways using grid technique, if for any reason at all, ' because it's there.' A third category of chapters appears based on the premise that if grid techniques exist they must be used and all that remains is the discovery of a problem to apply them to (i.e., Kelly's theory is incidental), whilst there are other examples where authors appear to wish simply to show that grids can be given to anyone, with a little ingenuity. One would not, in short, be surprised to find a chapter called ' Grid Techniques applied to blind deaf mutes in the iron lung,' and this raises precisely the issue of what advance of knowledge is represented by chapters in this latter category.

In short, the theoretical impact of the book is hard to assess since so many of the issues it tackles appear to be *simply* examples of where the technique can be used, rather than comparisons of this *versus* that technique. Presumably any technique can be applied almost anywhere once the relationship of technique to theory is weakened.

The theoretical assessment of the book is further complicated by the fact that it bears all the signs of having been long in the pipeline (written about 1972/1973 by the look of it) with later references being very few and largely confined to authors' own work. This is particularly important since those chapters that do make theoretically interesting statements have occasionally been overtaken by events when they have identified areas that need to be developed.

One would certainly like to see a good book answering the challenge of Bannister and Fransella (1971) in showing the many new insights given to old problems by use of Kelly's theory and associated methods. Whilst this present book undoubtedly will be a useful ' workshop manual ' to those who use grid technique, it is not clear that it represents more than this.

STEVE DUCK.

STUBBS, M., and DELAMONT, S. (Eds.) (1976). *Explorations in Classroom Observation.* London: Wiley, pp. 221, £6·50, $13·50.

This book presents a collection of nine papers concerned with ways of observing and thinking about the classroom world. The introductory paper by Delamont and Hamilton reviews the range of investigative methods in the field, justifying the study of classroom events and challenging the restrictions that the wide currency of the Flanders and other systematic observation procedures have placed upon the perceptiveness and sensitivity of classroom observers. Thus, along with systematic observational methods, Delamont and Hamilton offer descriptions and examples of the rich yield to be gained from anthropological-style observation in classrooms. Each of the remaining eight papers, which are presented in three groups, offers its distinctive lens to the classroom observer by inviting him to observe classroom life in a distinctive way and through the use of particular kinds of concepts.

Three papers deal with pupils' perceptions of school, exploring in particular the relationships between pupils' interactions and their shared perceptions of interpersonal and specifically pedagogical behaviour in school.

Four papers present approaches to the study of classroom interactions and classroom talk. The paper by Delamont discusses the fruitfulness of combining standard, structured observational methods with observational insights sensitive to and a rising out of a particular research setting. This writer, applying these two observational methods in a study of teachers' self-presentation in a girls' fee-paying school, reads engagingly like a new Jane Austen for the feminine classroom world. In their paper called ' Strawberries,' Walker and Adelman illuminate the importance of probing the established meanings of the verbal shorthand so often heard in classrooms and so incomprehensible to the newcomer, and they plead for the long-term study of classroom events and relationships. The papers by Stubbs and Torode offer the reader perspectives and descriptions for making different kinds of sense of classroom talk, Stubbs writing within a sociolinguistic framework, and Torode building a phenomenological account which goes beyond the teacher's apparent performance and probes the underlying controlling functions of the teacher's classroom discourse.

The final section of the book contains a study by Hamilton which combines classroom observation with other research techniques and, by attaching its observational insights to curriculum theory as well as to relevant school variables, illustrates how observational methods may be embedded within wider theoretical and organisational perspectives.

While many of the papers in this book report on researches or aspects of researches written about elsewhere, this particular collection is enjoyable as an effective invitation to the new reader to construe the classroom world in a multidimensional fashion. The verbal exchanges which are the cumbersome raw data of this field are, on the whole, typographically well and consistently handled. It is a pity, however, that the editors' introduction at the beginning of each paper are distinguished from the papers themselves in a variable and, I found, rather confusing way. This quibble apart, the book is to be welcomed as an attempt to present and to explicate some of the many possible approaches to the fascination of classroom life.

M. CAMERON-JONES.

TAYLOR, J. H. (1976). *The Half-Way Generation: A Study of Asian Youths in Newcastle-upon-Tyne.* London: NFER, pp. 267, £5·28.

This is an interesting, informative and readable account of a study in depth of the Indian and Pakistani boys (67 in all) who reached school leaving age in Newcastle over the six years 1962-67. As an SSRC Research Fellow at Newcastle University Mr. Taylor undertook the initial fieldwork for his study in 1968 (the year of some of the most virulent of Powell's anti-immigrant speeches) and followed up the young men until 1971. Only 5 of the 67 had been born in this country. Twenty-eight had come here either at primary school level or earlier and the remainder at 13 or over. Thus, the majority were child migrants. Hence the title of the book: *The Half-Way Generation.* The smallness of the group enabled Mr. Taylor to keep in touch with each individual and to do most of his interviewing in their homes. He quotes frequently from these interviews, and the frankness of the young men's replies (they were between 15 and

21 when interviewed) testifies to the degree of rapport he established with them. Unlike many studies published in recent years, this one is not primarily concerned with the educational attainments and employment of immigrants but with such questions as how far there is conflict between young Asians and their parents about the traditional values and norms of the parental culture, about how far the young people feel they belong to this country or identify with their parents' homeland, and to what extent they have succeeded in adjusting themselves to the culture, the economic situation and the social climate of this country.

One striking finding about educational performance does, however, emerge. Matched against a control group of English youths from the same schools as they had attended, the Asian youths came out better in terms of the proportions continuing full-time education after 15, the proportions still in full-time education when interviewed, and the proportions who gained admission to higher education. On certificates gained by school leavers the two groups were evenly matched. Mr. Taylor accounts for the superior performance of the Asians in terms of the distinctive caste backgrounds of many of the parents, the high value placed on education in Indo-Pakistani culture, the supportive warmth of the Asian family and, possibly also, though this factor did not come through strikingly in the interviews, the realisation that to succeed in a predominantly white society, the coloured person needs to aim higher than his white counterparts.

The first part of the book sets the youths in the context of the cultural backgrounds of their parents (almost 75 per cent had come from the Indian or Pakistani part of the Punjab), the city of Newcastle and the Elswick area of the city where most of them had settled, their fathers' jobs and the prevailing racial climate. Thirty-seven (57 per cent) of the fathers were self-employed (credit drapers, shopkeepers, etc.).

Interesting though these aspects of the context are, it is those parts of the book which deal with the young people themselves, particularly the chapters on *religion, friendship and marriage*, *a feeling of not belonging*, *a question of identity*, and *measure of adjustment*, that are most likely to interest educational psychologists and teachers. To summarise these in a brief review is not possible, but here are quotations from two of the young men themselves. They emphasise how strong is the feeling of many young Asians that they do not belong to England and how aware they are of the ambiguity of their identity. The first is from Lal, a Kenyan, who had spent seven years in this country: " No matter how western, how high your standard of living, you're still coloured, and they try to cast you as inferior or they try to prove you're inferior." The second is from Surinder, who was born in this country but had experienced hostility at school: " We're what you might call unpredictable Indians. We don't know where we stand."

A. LAING.

TIZARD, J., MOSS, P., and PERRY, J. (1976). *All Our Children: Pre-School Services in a Changing Society.* London: Temple Smith/New Society, pp. 252, p. £2·25, c. £5·00.

In recent years widespread interest has been shown in the care and education children prior to the age of compulsory schooling. Much controversy has been generated over issues such as the extent to which the state should provide services for these children, the nature of these services and the most appropriate ways of organising them. Differing opinions, too, have been expressed about the role of professionals in the care of young children, parental involvement in pre-school education and the philosophy which should guide teaching in the nursery school. In the course of the debate, it is not surprising if sweeping generalisations have often been made, not least on the subject of nursery education, seen by some as the panacea for the ills of society, by others as offering little to the children in most need of stimulation.

This well-balanced book by Jack Tizard and his associates provides a corrective to the more extreme views, both positive and negative, on the value of providing services for pre-school children. While the writers play down the long-term effects of nursery schooling, emphasising that goals at this stage should be modest and immediate rather than concerned with hoped-for benefits in the distant future, they point the way to giving many young children a happier and more stimulating start to life than they now enjoy.

The first four chapters portray in some detail the historical background to our present ' inadequate and irrationally organised ' services. Current provision for the under-fives in England and Wales is then compared (mainly unfavourably) with that existing in several other West European countries. The growing social pressures on mothers and families are highlighted, and two chapters are devoted to working mothers, whose problems, in the writers' view, have been neglected in the past by both governments and researchers. Certainly, the high incidence of disturbance reported among working mothers with young children gives cause for concern. In the concluding sections, Tizard and colleagues deal with a variety of questions relating to the provision of pre-school services, offering guidelines for the development of a comprehensive and effective pre-school service. Their discussion of the needs of disadvantaged children in the early years contains many points of interest, though it suffers from over-conciseness, the studies drawn from being highly selective.

It is a pity, perhaps, that after recognising in the preface how little attention has hitherto been given to the role of the father, Tizard and his co-authors do nothing to remedy the situation in this book. It is a pity, too, that little, if anything, is said about the needs of young children with mental, physical or emotional handicaps, for whom adequate pre-school services are especially important. Even so, it is very valuable, at this stage in the development of pre-school provision, to have a book which takes a broad view and gives the reader an overall perspective covering the different forms of provision for young children up to the age of 5. Praiseworthy, too, is the stress on giving parents what they themselves want rather than what they are said to need by other people and official agencies. The authors state their own position firmly but with due caution where this is called for. Although their arguments are supported by a mass of facts and figures, the book remains lucid and readable throughout.

This book can be highly recommended to all concerned with young children. Even if it contains little new material, students will find it an extremely useful overview and it should be compulsory reading for those responsible for guiding policy in the development of pre-school services. The book should confirm that there is a strong demand for these services, but also encourage policy-makers to remember that the debate is about quality as well as quantity.

MAURICE CHAZAN.

VINCENT, D., and CRESSWELL, M. (1976). *Reading Tests in the Classroom.* Slough: NFER, pp. 185, n.p.

This book contains a large amount of good sense about the value and limitations of reading tests in a classroom situation. The authors are careful to avoid any suggestion of a ' mystique,' and are at pains to point out that a test is not invalidated because unskilled people misuse it. There is also some helpful guidance on statistical matters, and some very useful discussion of different types of test. I particularly enjoyed the excellent section on ' criterion referenced testing.' Briefly, such testing checks whether certain skills have been achieved rather than attempting to compare a child with some ' norm.' If this concept is taken seriously teachers will be encouraged to consider what are realistic ' targets ' for particular children and there will be fewer invidious comparisons between one child and another; more sensibly, each child will be compared with *himself.*

The difficulties inherent in the notion of a ' true ' reading age are perhaps passed over somewhat lightly. (If the concept is a confused one, then placing inverted commas round ' true ' will not remedy the situation!) It was also a little disappointing that no mention was made of the Gillingham-Childs Phonic Proficiency Scales, and that Vernon (1957) is referred to in preference to Vernon (1971). The emphasis on reading skills as opposed to other kinds of language skill is, of course, deliberate, but I hope that this will not encourage teachers to concentrate too much on reading at the expense of spelling; in the experience of myself and colleagues this has particularly unfortunate effects in the case of children who display problems in certain specific areas of learning.

All in all, however, this is a very sensible book.

T. R. MILES.

WEST, C. K., and FOSTER, S. F. (1976). *The Psychology of Human Learning and Instruction in Education.* Belmont, California: Wadsworth, pp. xiii+278, n.p.

There is no shortage of educational psychology books. When asked to review yet another the heart sinks. It sinks further when a brief glance at the contents reveals the familiar topics of theories of learning, motivation and reinforcement, retention and transfer, thinking, problem solving, creativity, individual differences, attitude learning and measurement and evaluation. The heart settles on rock bottom when the authors, in common with a long line of illustrious predecessors, claim to " provide a psychological rationale for decisions about teaching and learning " and to equip the student with those concepts necessary for further reading in the psychological literature on instruction. Despite the familiarity of content and purpose, however, West and Foster's book has a unique quality of emphasis and sufficiently contemporary focus to commend it as a useful contribution to a college of education library.

Its unique emphasis is that it is very much teacher-task centred. Through all the chapters the stress is on what a teacher needs to think about, do and evaluate. In chapter 1, the foundations of the book are laid on a five-stage model of teaching on which subsequent ideas are to be constructed. The five stages involve the teacher in a consideration of (1) the setting of objectives, (2) the assessment of pupils' capabilities, (3) the selection and organisation of experiences, (4) the guidance of learning and (5) the evaluation of learning.

In chapter 2, classical and modern theories of learning and instruction are surveyed briefly and evaluated in terms of their contribution to teachers' thoughts and actions with respect to the five stages of instruction. In chapter 3 the pupil's ' frame of reference ' is introduced. Here, differences between pupils in terms of their concepts, needs, interests, values and cognitive structures are discussed. The influence of these factors on perception and cognition and their pertinence to the design of the five stages of teaching are explored. There is a return to the problem of differences between individual learners in chapter 7. This is a very useful chapter in which once again the emphasis is on constructive response rather than pathological dismay. There is a discussion of mastery learning, programmed instruction and open education as solutions to the problems arising out of different rates of learning and levels of interest amongst pupils. Other chapters deal very briefly with motivation, imitation, transfer of learning, thinking (including cognitive development), problem solving, creativity and evaluation of learning.

Throughout the book competing, formal, psychological theories are discussed on each issue and at the end of each chapter the pertinence of these issues for the instructional model is considered. The last chapter summarises these implications in a final overview. The emphasis on the teacher and her task is reflected in what is omitted as well as in what is treated. There is no discussion of backward children, gifted children, delinquent children, working-class children, myopic children nor any other pseudo-pathological children and the authors are to be congratulated on this. Of course, 'The Child' is not missing from the book since the discussion of the learner's frame of reference permeates every chapter.

Despite its qualities the book is unlikely to be useful as a general-issue text. The treatment of all the psychological material is extremely brief and frequently unnecessarily complex. For example, five theories of transfers are dealt with in two pages and the notions of relativism and frame of reference are introduced via a discussion of modern physics. The chapter on memory is particularly heavy going; the authors rush too soon into the concerns of psychologists rather than educators. To be fair, West and Foster claim that their book will challenge most students and that background study in psychology would be useful. They are quite right.

There is much that irritates in the book. It is claimed that reinforcement "produces acquisitions"; on p. 58 there is an unnecessarily obscure discussion of relativism after which the authors distinguish between 'events' and 'phenomena' and then they misuse the distinction on p. 59; on p. 64 there seems to be some confusion in their minds between 'rote' as a process and 'rote' referring to materials to be learned and on the same page the concept of 'bit' (of information) is introduced with no explanation nor entry in the glossary; p. 69 contains a discussion on anxiety which is confusing, contradictory and inconsequential; on p. 74 Hernandez Peon's study of 1956 is deployed to sustain a point about needs and perceptions (i.e., despite the criticisms made of this study) and on p. 76 it is claimed that only recently have pupils' interests been taken into account in planning instruction—a claim that William James and Alexander Bain would have found dated.

But perhaps the major irritation arises out of the apparent ease with which the authors shift from psychological description to educational prescription. Frequently the shift is merely inconsequential (for example following the discussion on anxiety) but at times the psychological preliminaries seem tortuously contrived. After their overview of motivation and reinforcement, for example, it is claimed that teachers should use competition and co-operation and that when grouping children they should use small groups, large groups and one-to-one arrangements. Now it is not disputed that there are sensible arrangements to make from time to time in classrooms. What is in dispute is the manner in which psychology is used to 'illuminate' or 'justify' such decisions. In this respect the book is no better and no worse than the vast majority of educational psychology books.

Where the authors emphasise psychology they are unconvincing; where they emphasise the teacher's task they are generally impressive. Which leaves their book rather better than average.

CHARLES DESFORGES.

WOLFF, J. G. (1973). *Language, Brain and Hearing: An Introduction to the Psychology of Language with a Section on Deaf Children's Learning of Language.* London: Methuen, pp. xiv+237, £3·20.

As its subtitle indicates, this book is in two sections. The first, and larger, part is intended as an introduction to psycholinguistics for students of psychology and linguistics, educational psychologists, speech therapists and teachers of the deaf. This section is well and lucidly written; Mr. Wolff clearly has a talent for explaining ideas in this area of psychology.

In the second part of the book, the author attempts to solve the knotty problem of teaching language to hearing-impaired children. On p. 153, the most important frequencies for speech perception are said to be between 500 and 200 Hz, surely a misprint for 2000 Hz. In any case so narrow a range is questionable in relation to children born deaf and hence lacking a knowledge of the structural redundancies of speech and language. It is odd that Mr. Wolff should fail to make this point after so persuasively explaining the role of redundancy in language learning in the first half of his book. On the following page the book states that "children with moderate or severe losses (of hearing) stand a fair to poor chance of developing speech and language if they have hearing aids and amplification." The author surely means if they *do not* have hearing aids and amplification. In fact, most children in the moderate range in this country will be attending either normal schools or partially hearing units and will acquire near normal speech through amplification. Moreover, many of those in the severe partially hearing group (losses 60–85 dB) will develop speech perfectly intelligible to strangers given, of course, early, regular and efficient use of aided hearing together with experience of natural linguistic communication from which to infer the vocabulary and structure of English.

Mr. Wolff believes that the best strategy in dealing with the problem of pre-lingual hearing impairment is to begin with a consideration of the extreme or ' pure ' case of complete deafness; " if we can find a solution to this problem, this is bound to help us solve the less severe problems." The solution lies, in his view, in finding some " substitute primary medium " of communication (SPM) to take the place of, or act in support of, spoken language. The author's logic would be impeccable if a satisfactory SPM could be found. Much depends on the efficiency of any proposed SPM relative to that of the admittedly attenuated information available through lip-reading and amplified hearing. The danger of the argument is that, in a flush of theoretical enthusiasm, an SPM will be introduced for the whole range of hearing-impairments including " the less severe problems " and will prove even less efficient for the majority of children than that currently available.

The next step in Mr. Wolff's argument is to consider in turn a series of candidates for SPM. The use of lip-reading and amplified residual hearing is given short shrift in a discussion which, it must be said, reveals Mr. Wolff's ignorance about modern methods in the education of the deaf. After considering a potpourri of other alternatives, the author arrives at his own solution. This is to ' augment ' Dr. Orin Cornett's Cued Speech (a system modestly intended to resolve some of the ambiguities of lip-read patterns by the use of the fingers as cues) so that *every* phoneme is recognisable from the fingers alone. The author does not consider the problem of timing from which his system suffers and its advantages over simple finger-spelling are hard to detect. It is notable that Mr. Wolff's idea, first published in 1971, has not been taken seriously even in the present climate of disenchantment with oral methods which is evident in some quarters. Mr. Wolff has made a brave try, but it seems unlikely that anyone can arrive at an armchair solution to this intractable problem which has perplexed and divided teachers of the deaf for two centuries.

The first section of this book can be recommended to its intended readership but the second half must be taken with a large pinch of salt.

W. D. HINE.

WOODHEAD, M. (1976). *Intervening in Disadvantage: A Challenge for Nursery Education.* Slough: NFER, pp. 129, £2·15.

This review of British research into pre-school education for disadvantaged children is timely in that it puts in perspective what has been achieved in this country to date. So many references have been made recently to the ' failure ' of compensatory education programmes that one might be forgiven for thinking that a negative view of the value of early intervention is based on a substantial body of evidence. This is certainly not the case as far as Britain is concerned, since there have been few pre-school projects for disadvantaged children here, and these have been relatively limited in scale of operation. While Martin Woodhead's book does little to encourage optimism, it does show that a review of British pre-school action-research is largely confined to the EPA projects (directed by A. H. Halsey and sponsored jointly by the Department of Education and Science and the Social Science Research Council), the NFER's own ' Pre-School Project ' (directed by H. L. Williams), and Joan Tough's work on young children's language development (supported by the Schools Council, but not concerned mainly with disadvantaged children).

The book opens with an introductory chapter on experiments in nursery education, in which an all-too-brief reference is made to the American compensatory projects. As, in the author's own words, the work carried out in the USA has greatly influenced British thinking and is the direct forerunner of much of the research in this country, a longer and more searching look at the American scene would have been relevant here. There is little doubt that we can learn much from the American experience.

The core of the review is concerned with the development and evaluation of pre-school curricula for disadvantaged children. Woodhead sketches the background to current controversies vividly, underlining the basic conflict between the more 'structured' approaches of compensatory programmes and traditional nursery education in this country. He discusses and evaluates the intervention programmes carried out at the nursery stage, making comparisons between the EPA and NFER projects. These comparisons are skilfully made, though the constant switching from one project to another may result in some confusion for the reader. Woodhead stresses that the commitment of staff to the aims and methods of a programme may be more crucial to its success than the particular approach used. He also underlines the need to look at the value of nursery education for disadvantaged children from the short-term rather than the long-term point of view. In the present state of our knowledge, this is probably right, but, as previously stated, there is too little research evidence to justify any conclusions, in the British context, about the long-term effects of early educational intervention. In spite of all the difficulties involved in long-term research, we need studies in depth of the effects of providing sustained help and support for disadvantaged children throughout their school careers.

A chapter is devoted to work with parents of pre-school children, an area which has received little attention hitherto, but in which a promising start has been made. Finally, Woodhead reviews recent research and developments in the provision of services for pre-school children, suggesting ways in which these might be improved. He argues for a variety of patterns of nursery provision to meet the needs of families living in different conditions, and pleads for a policy of actively seeking out children in particular need of nursery education: the limited evidence available suggests that, at present, it is the more advantaged families which are most likely to benefit from pre-school provision.

Overall, the book is characterised by a balanced and cautious approach to the questions discussed. The author is well-informed and up-to-date, and has succeeded in producing a readable and jargon-free review of research. The book will be found valuable by students and can also be recommended to teachers concerned with disadvantage, who must by now be thoroughly confused about the direction in which they ought to go. They should emerge from reading the book with a greater understanding of the complex issues involved, even if they are likely to remain confused about their own role in helping disadvantaged children.

MAURICE CHAZAN.

PUBLICATIONS RECEIVED

Books for review have been selected from the following. (" n.p." means " no price stated.")

ANTHONY, E. J., and GILPIN, D. C. (1977). *Three Clinical Faces of Childhood.* New York: Spectrum Publications, Distributor John Wiley, pp. 255, £11·25, $19·00.

ARGYRIS, C. (1976). *Increasing Leadership Effectiveness.* New York: Wiley, pp. xvi+286, £11·90, $19·00.

AULT, R. L. (1977). *Children's Cognitive Development: Piaget's Theory and the Process Approach.* New York: Oxford University Press, pp. ix+193, £2·75.

BASSIN, A., BRATTER, T. E., and RACHIN, R. L. (Eds.) (1976). *The Reality Therapy Reader: A Survey of the Work of William Glasser, M.D.* New York and London: Harper and Row, Ltd., pp. xi+691, £11·25.

BECKER, W. C., ENGELMANN, S., and THOMAS, D. R. (1975). *Teaching* 1: *Classroom Management.* Chicago: Science Research Associates, pp. viii+359, £4·85.

BECKER, W. C., ENGELMANN, S., and THOMAS, D. R. (1976). *Teaching* 2: *Cognitive Learning and Instruction.* Chicago: Science Research Associates, pp. viii+263, £4·80.

BERG, L. (1977). *Reading and Loving.* London: Routledge and Kegan Paul, pp. 136, p. £1·20, c. £2·75.

BERGLING, K. (1976). *The Development of Hypothetico-Deductive Thinking in Children.* Stockholm: Almqvist and Wiksell, New York and London: John Wiley, pp. xix+127, £7·50, $12·10.

BERTHOUD, R. (1976). *The Disadvantages of Inequality.* London: Macdonald and Janes, pp. 207, £4·95.

BINDER, L. (Ed.) (1976). *The Study of the Middle East: Research and Scholarship in the Humanities and Social Sciences.* New York: John Wiley, pp. 648, £11·00, $20·35.

BOLTON, N. (1976). *The Psychology of Thinking.* London: Methuen, pp. ix+291, £2·75.

BRODY, E. B., and BRODY, N. (1976). *Intelligence: Nature, Determinants, and Consequences.* New York: Academic Press, pp. x+241, £8·85, $12·50.

BROWN, C. (1976). *Australian Reports on Education.* Victoria, Australia: ACER, pp. 47, $3·50.

BROWN, F. G. (1976). *Principles of Educational and Psychological Testing* (2nd edit.). New York: Holt, Rinehart and Winston, pp. viii+504, £9·50.

BROWN, G. (1977). *Child Development.* London: Open Books, pp. viii+167, £1·95.

BROWNMILLER, S. (1977). *Against Our Will: Men, Women and Rape.* Harmondsworth: Penguin, pp. 472, £0·95.

BURGESS, T. (1977). *Education After School.* Harmondsworth: Penguin, pp. 256, £0·80.

BUSS, A. R., and POLEY, W. (1977). *Individual Differences : Traits and Factors.* New York : Gardner Press, Distributor John Wiley, pp. ix+275, £11·25, $19·00.

CLEVERLEY, J., and JONES, P. (1976). *Australia and International Education.* Victoria, Australia: A.C.E.R., pp. 45, $2·25.

CORTIS, G. (1977). *The Social Context of Teaching.* London: Open Books, pp. x+169, £1·95.

COTTLE, T. J. (1976). *Perceiving Time: A Psychological Investigation with Men and Women.* New York: John Wiley, pp. xiv+267, £10·60, $17·20.

DAVIES, A. (Ed.) (1977). *Language and Learning in Early Childhood.* London: Heinemann, pp. vii+222, £4·75.

DENNIS, W., and DENNIS, M. W. (Eds.) (1976). *The Intellectually Gifted : An Overview.* New York: Grune and Stratton, pp. xiii+336, £15·95, $22·50.

DOWNING, J. (1976). *Learning to Read with Understanding.* Cape Town: Juta, pp. vii+101, n.p.

DUFF, T. S., and CLARK, M. L. (1976). *Listening in the Primary School.* Victoria, Australia: A.C.E.R., pp. xiv+208, $8·50.

ELKIND, D. (1977). *Child Development and Education : A Piagetian Perspective.* New York: Oxford University Press, pp. xiii+274, p. £3·75, c. £8·95.

Estes, W. K. (Ed.) (1976). *Handbook of Learning and Cognitive Processes, Volume* 4—*Attention and Memory.* Hillsdale, New Jersey: Lawrence Erlbaum Associates, Distributor John Wiley, pp. x+436, £15·90, $25·00.

Eysenck, H. J. (1976). *The Measurement of Personality.* Lancaster: MTP Press, pp. xviii+511, £12·50.

Eysenck, H. J., and Eysenck, S. B. G. (1977). *Psychoticism as a Dimension of Personality.* London: Hodder and Stoughton, pp. xi+232, £6·50.

Fitzgerald, M. (1977). *Prisoners in Revolt.* Harmondsworh: Penguin, pp. 278, £1·00.

Fitzherbert, K. (1977). *Child Care Services and the Teacher.* London: Temple Smith, pp. x+225, £4·65.

Floyd, M. K. (1976). *Abortion Bibliography for* 1975. New York: Whitston Publishing Co., pp. xvii+295, $12·00.

Fontana, D. (1977). *Personality and Education.* London: Open Books, pp. viii+178, £1·95.

Fraisse, P. (Ed.) (1976). *L'Année Psychologique* 1976, 2. *Paris*: Presses Universitaires de France, pp. 354, n.p.

Freud, A. (Ed.) (1977). *The Psychoanalytic Study of the Child* XXXI. London: The Hogarth Press, pp. vii+573, £12·00.

Frostig, M. (1976). *Education for Dignity.* New York: Grune and Stratton, pp. xiii+208, £7·95, $13·00.

Gath, D., Cooper, B., Gattoni, F., and Rockett, D. (1976). *Child Guidance and Delinquency in a London Borough.* London: Oxford University Press, pp. xiii+190 £7·00.

Getzels, J. W., and Csikszentmihalyi, M. (1976). *The Creative Vision: A Longitudinal Study of Problem Finding in Art.* New York: John Wiley, pp. x+293, £14·30, $22·80.

Gibbons, J. D. (1976). *Nonparametric Methods for Quantitative Analysis.* New York: Holt, Rinehart and Winston, pp. xv+463, £9·75.

Gibson, D., and Brown, R. I. (Eds.) (1976). *Managing the Severely Retarded: A Sampler.* Springfield, Illinois: Charles C. Thomas, pp. xvi+482, $22·75.

Gruijter, D. N. M. de, and Kamp, L. J. T. van der (Eds.) (1976). *Advances in Psychological and Educational Measurement.* London: John Wiley, pp. xi+320, £14·00, $29·00.

Hammersley, M., and Woods, P. (Eds.) (1976). *The Process of Schooling : A Sociological Reader.* London and Henley: Routledge and Kegan Paul, pp. vi+232, p. £1·90, c. £3·75.

Harper, A. E., and Misra, V. S. (1976). *Research on Examination in India.* New Delhi: National Council of Educational Research and Training, pp. xxiv+318, n.p.

Harvey, J. H., Ickes, W. J., and Kidd, R. F. (Eds.) (1976). *New Directions in Attribution Research, Volume I.* Hillsdale, New Jersey: Lawrence Erlbaum Associates, Distributor John Wiley, pp. x+467, £15·90, $24·95.

Hawkins, P. R. (1977). *Social Class, The Nominal Group and Verbal Strategies.* London, Henley and Boston: Routlege and Kegan Paul, pp. xii+242, £7·50.

Herriot, P. (1976). *An Introduction to the Psychology of Language.* London: Methuen, New York: Barnes and Noble, pp. 197, £2·10.

Illich, I., and Verne, E. (1976). *Imprisoned in the Global Classroom.* London: Writers and Readers Co-operative, pp. 64, £0·45.

Janov, A. (1977). *The Feeling Child.* London: Abacus, Sphere Books, pp. 286, £1·50.

Katz., J., and Hartnett, R. T. (1976). *Scholars in the Making: The Development of Graduate and Professional Students.* Cambridge, Mass.: Ballinger Publishing Co., Distributor John Wiley, pp.xv+287, £10·65, $16·50.

Kohl, H. R. (1977). *Writing, Maths. and Games in the Open Classroom.* London: Methuen, pp. xi+200, £2·35, c. £4·00.

Laing, R. D. (1976). *The Facts of Life.* London: Allen Lane, pp. 143, £3·25.

Lancaster, J., and Gaunt, J. (1976). *Developments in Early Childhood Education.* London: Open Books, pp. ix+145, p. £1·75, c. £3·90.

Laslett, R. (1977). *Educating Maladjusted Children.* London: Crosby Lockwood Staples, pp. x+269, £7·50.

Leach, D. J., and Raybould, E. C. (1977). *Learning and Behaviour Difficulties in School.* London: Open Books, pp. viii+168, £1·95.

Lee, G. W. (1977). *Sex Education and the Mentally Retarded.* London: N.S.M.H.C., pp. 14, £0·50.

Lee, P. C., and Stewart, R. S. (1977). *Sex Differences: Cultural and Developmental Dimensions.* New York: Urizen Books, pp. 478, p. £3·90, c. £9·60.

Lee, T. (1976). *Psychology and the Environment.* London: Methuen, pp. 143, p. £0·80, c. £2·50.

Lidz, T. (1976). *The Person* (revised edit.). New York: Basic Books, pp. xxiii+615, £8·20.

Lunt, H. N. (Ed.) (1976). *Language and Language Teaching: Current Research in Britain* 1972-75. London: Longman, pp. xi+259, n.p.

Luria, A. R. (1977). *Cognitive Development: Its Cultural and Social Foundations.* Cambridge, Mass., and London: Harvard University Press, pp. xvi+175, £9·40.

Macintosh, H. G. (1977). *Assessing Attainment in the Classroom.* London: Hodder and Stoughton, pp. 47, £0·80.

Mann, L., and Sabarino, D. A. (Eds.) (1976). *The Third Review of Special Education.* New York: Grune and Stratton, pp. viii+277, £16·25, $25·00.

Martin, F. M., and Murray, K. (Eds.) (1976). *Children's Hearings.* Edinburgh: Scottish Academic Press, pp. xv+242, p. £3·75, c. £5·00.

McClelland, D. C., Atkinson, J. W., Clark, R. A., and Lowell E. L. (1976). *The Achievement Motive.* New York: Irvington Publishers, Distributor John Wiley, pp. xxii+386, £11·90, $19·00.

Miller, E. (1977). *Abnormal Ageing: The Psychology of Senile and Presenile Dementia.* London: John Wiley, pp. x+166, £5·75, $11·00.

Mitford, J. (1977). *The American Prison Business.* Harmondsworth : Penguin, pp. 329, £1·00.

Morrison, D. W. (1977). *Personal Problem Solving in the Classroom: The Reality Technique.* New York: John Wiley, pp. xii+186, £4·50, $7·55.

Morton, J., and Marshall, J. C. (Eds.) (1977). *Psycholinguistics* 1: *Developmental and Pathological.* London: Elek Science, pp. vii+160, £4·95.

National Foundation for Educational Research in England and Wales (1976). *Register of Educational Research in the United Kingdom* 1973-1976. Slough: NFER, pp. 368, £12·00.

Neligan, G. A., Kolvin, I., Scott, D. McL., and Garside, R. F. (1977). *Born Too Soon or Born Too Small.* London: Spastics International Medical Publications, pp. xx+101, £4·50.

Paterson, B. (1977). *Help!* Harmondsworth: Penguin, pp. 378, £1·00.

Patrick, R. J. (1976). *An Annotated Bibliography of Recent Continuing Education Literature.* Stanford, Calif.: The ERIC Clearing house on Information Resources, pp. 52, $5·00.

Postic, M. (1977). *Observation et Formation des Enseignants.* Paris : Presses Universitaires de France, pp. 336, n.p.

Poteet, G. H. (1976). *Death and Dying: A Bibliography* 1950-1974. New York: Whitston Publishing, pp. xxxvii+192, $12·50.

Reese, H. W. (1976). *Basic Learning Processes in Childhood.* New York: Holt, Rinehart and Winston, pp. xi+174, £2·75.

Riding, R. J. (1977). *School Learning: Mechanisms and Processes.* London: Open Books, pp. viii+167, £1·95.

Ross, K. N. (1976). *Searching for Uncertainty.* Victoria, Australia: A.C.E.R., pp. xii+56, $3·00.

Rothenberg, A., and Hausman, C. R. (Eds.) (1976). *The Creativity Question.* Durham, N. Carolina: Duke University Press, pp. xiv+366, p. $7·95, c. $14·75.

Rowan, J. (1976). *Psychological Aspects of Society, Book* 3: *The Power of the Group.* London: Davis-Poynter, pp. 212, £5·00.

Rowerton, W. E. (Ed.) (1976). *Revitalising Educational Psychology : Readings in Methods and Substance.* Chicago, Illinois: Nelson-Hall, pp. 390, $15·00.

Rutter, M., and Madge, N. (1976). *Cycles of Disadvantage*. London: Heinemann, pp. vi+413, p. £2·50, c. £6·50.

Sharpe, S. (1976). *'Just Like a Girl': How Girls Learn to be Women*. Harmondsworth: Penguin, pp. 328, £0·95.

Shennan, V. (1976). *Help Your Child to Understand Sex*. London: N.S.M.H.C., pp. 7, £0·30.

Spielberger, C. D., and Diaz-Guerrero, R. (1976). *Cross-Cultural Anxiety*. Washington, D.C.: Hemisphere Publishing Corporation, Distributor John Wiley, pp. xii+195, £14·30, $23·00.

Spooncer, F. A. (1977). *The Development of Reading*. London: Hodder and Stoughton, pp. 31, £0·60.

Spooncer, F. A. (1977). *The Evaluation of Reading*. London: Hodder and Stoughton, pp. 31, £0·60.

Stone, L. (Ed.) (1977). *Schooling and Society: Studies in the History of Education*. Baltimore and London: John Hopkins University Press, pp. xvii+263, £11·20.

Tompkins, M., and Shirley, N. (1976). *Serials in Psychology and Allied Fields*. New York: Whitston Publishing Co., pp. xi+472, $22·50.

Travers, J. F. (1977). *The Growing Child: Introduction to Child Development*. New York: John Wiley, pp. ix+531, £8·00, $13·75.

Triche, C. W., and Triche, D. S. (1976). *The Euthanasia Controversy* 1812-1974. New York: Whitston Publishing Co., pp. viii+242, $18·00.

Triche, C. W., and Triche, D. S. (1976). *Drug Abuse Bibliography for* 1974. New York: Whitston Publishing Co., pp. v+450, $18·00.

Tuck, M. (1976). *How Do We Choose*? London: Methuen, pp. 144, p. £0·80, c. c. £2·50.

Tudor-Davies, E. R. (1977). *Parents Information Bulletin* 1977 *Supplement*. London: N.S.M.H.C., pp. 25, £0·20.

Turgenev, I. (1977). *First Love*. Harmondsworth: Penguin, pp. 93, £0·45.

West, C. K., and Foster, S. F. (1976). *The Psychology of Human Learning and Instruction in Education*. Belmont, Calif.: Wadsworth, pp. xiii+302, n.p.

Williams, J. (1977). *The Impact of Research*. London: L.V.C.E.R.I., pp. 7, £0·15.

Wirth, J. E. (Ed.) (1976). *Assessing Linguistic Arguments*. Washington, D.C.: Hemisphere Publishing, Distributor John Wiley, pp. viii+280, £11·90, $19·00.

Woodhead, M. (Ed.). *An Experiment in Nursery Education*. Slough: NFER, pp. 123, £3·95.

PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

available from the Society's offices

St. Andrews House, 48 Princess Road East, Leicester, LE1 7DR

Careers in Psychology	50p
Psychology and Psychologists: A Pamphlet for Schools	15p
Teaching Educational Psychology in Training Colleges	25p
Technical Recommendations for Psychological and Educational Tests	10p
Aspects of Autism. Edited by P. J. Mittler	50p
The Psychological Study of Deafness and Hearing Impairment. Edited by M. Rodda	40p
Postgraduate Qualifications and Courses in Psychology 1976	40p

PRESIDENTIAL ADDRESSES

1965 Perceptual Defence and the Engineering Psychologist. Dr. Donald Broadbent	10p
1968 Dissonance and Stress in the Teaching of Psychology: Some Reflections of an Accidental Psychologist: Professor George Seth	10p
1969 Changing Horizons: An Essay in Autobiography: Dr. Boris Semeonoff	10p
1970 Choosing: Professor R. J. Audley	10p
1971 In Search of an Idiographic Psychology. By Professor H. Gwynne Jones	10p
1972 Psychology today and tomorrow. By Professor Harry Kay	10p
1973 Psychology in Society: Ends or End. By Professor Max Hamilton	10p
1974 On Taking Sides: Professor B. M. Foss	10p
1975 Thought and the Brain: Professor O. L. Zangwill	10p
1976 Psychology and Social Policy: Professor Jack Tizard	10p

C. S. MYERS LECTURES

Remembering Dr. Myers: Sir Frederic Bartlett	10p
Beyond Industrial Psychology: Dr. R. Conrad	10p
Practice at Liking: A Study in Experimental Aesthetics: Professor D. W. Harding	10p
C. S. Myers in Retrospect: Professor Alec Rodger	10p
On Applied Psychology in Britain: Professor G. C. Drew	10p
Measuring Psychiatric Morbidity in a South Indian Population: Professor G. M. Carstairs	10p

MEMORANDA OF EVIDENCE

To the Royal Commission on the Law Relating to Mental Illness and Mental Deficiency 1954-1957	10p
To the Royal Commission on the Penal System in England and Wales	10p
To the Royal Commission on Medical Education	10p
To the Ministry of Health and Department of Education and Science on the Responsibility for the Education of Mentally Handicapped Children	10p
To the Advisory Council on the Penal System (Sub-committee on Detention Centres)	10p
To the Departmental Committee on the Adoption of Children	10p
To the Department of Health and Social Security Sub-committee on the Role of Psychologists in the Health Services	10p
To the Butler Committee on the Law relating to the Mentally Abnormal Offender	10p
To the Warnock Committee on the Enquiry into Special Education (Summary)	10p

MISCELLANEOUS

Principles governing the employment of psychological tests and clinical instruments	7p
Psychological Tests: A Statement by the British Psychological Society	7p
Memorandum on the Establishment and Accommodation of Psychologists in General Hospitals	7p
The School Psychological Service (England and Wales)	7p
The changing role of the Educational Psychologist in Child Guidance Services in Scotland	10p
Teaching Psychology to Nurses (DCP Report)	10p
Training in Psychotherapy (DCP Report)	10p
Guidelines for the In-service Training of Probationer Clinical Psychologists	10p

(All prices on the above publications include postage)

INTERNATIONAL JOURNAL OF PSYCHOLOGY
JOURNAL INTERNATIONAL DE PSYCHOLOGIE

Published quarterly by the International Union of Psychological Science

The International Journal of Psychology is devoted to cross-cultural comparisons of psychological phenomena in general psychology, child psychology and social psychology. It also deals with international confrontation of theories and methods on which psychologists from various countries differ because of their native culture and training. Oriented towards basic research and theory, it publishes scientific papers in English or in French. In each issue, an International Platform gives news and informations on internationally significant problems.

Editor (Scientific papers)
J. Leroux
100, Tiensestraat
3000 Louvain, BELGIUM

Associate Editor (International Platform)
W. Holtzman
University of Texas, Austin, U.S.A.

1974 VOLUME 9 No. 3

CONTENTS **SOMMAIRE**

Topological and Euclidean spatial features noted by children : A cross-cultural study
GUSTAV JAHODA, JAN B. DEREGOWSKI and DURGANAND SINHA.... 159

Ecology, culture and psychological differentiation
JOHN W. BERRY and ROBERT C. ANNIS 173

Le développement de la conservation des longueurs de deux horizontales décalées : étude comparée d'enfants canadiens-français et rwandais
ADRIEN PINARD 195

A cross-cultural study of recognition memory for faces
J. W. SHEPHERD, J. B. DEREGOWSKI and H. D. ELLIS 205

Indian and Canadian choice behaviour in a Maximizing Difference Game and in a Game of Chicken
DAVID W. CARMENT 213

INTERNATIONAL PLATFORM FOR PSYCHOLOGISTS — TRIBUNE INTERNATIONALE DES PSYCHOLOGUES

Scientific and professional ethics and conduct for psychologists. A statement by the Executive Committee of I.U.P.S. 225

Code déontologique pour psychologues. Une déclaration du Comité Exécutif de l'I.U.P.S. 227

Continuing concerns in cross-cultural ethics : A report
J. L. TAPP, H. C. KELMAN, H. C. TRIANDIS, L. S. WRIGHTSMAN and G. V. COELHO 231

Annual subscription (4 issues) **$17.00**
Single issue **$5.00**

Should be sent to : **Centrale des Revues DUNOD-GAUTHIER-VILLARS**
24-26, Boulevard de l'Hôpital 75005 PARIS

EDUCATIONAL RESEARCH

Vol. 19, No. 2 — *Editor* : ALFRED YATES — *February*, 1977

R. B. NICODEMUS. Why science teachers adopt new curriculum projects.

R. W. FAIRBROTHER and J. R. L. SWAIN. The assessment of project work in Nuffield Advanced Biology and Nuffield Advanced Physical Science.

CAROL M. LOMAX. Interest in books and stories at nursery school.

JOHN GRAY. Teacher competence in reading tuition.

V. A. FURLONG and A. D. EDWARDS. Language in classroom interaction: theory and data.

MILENA M. JELINEK. Multi-racial education—3. Pupils' attitudes to the multiracial school.

F. J. O'HAGAN. Attitudes of offenders and non-offenders towards school.

BOOK REVIEWS.

The subscription is £4·75 a year (incl. postage); £6·50 Overseas (postage paid).

All orders should be sent to :

NFER Publishing Company Ltd. (Journals Department), 2 Jennings Buildings,
Thames Avenue, Windsor, Berks., SL4 1QS

The British Journal of Medical Psychology

Volume 49, Part 4, December, 1976

S. Stephanos, W. Biebl, and **F. G. Plaum.** Ambulatory analytical psychotherapy for the treatment of psychosomatic patients: A report on the method of ' relaxation analytique '

D. A. Shapiro. The effects of therapeutic conditions: Positive results revisited

Eric Bromley and **Leonard A. Lewis.** A factor analytic study of the Dynamic Personality Inventory using a psychiatric population

Krysia M. Yardley. Training in feminine skills in a male transsexual: A preoperative procedure

D. J. Sperlinger. Aspects of stability in the repertory grid

Horatio Fabrega, Jr. and **Stephen Tyma.** Language and cultural influences in the description of pain

Michael Herbert, Murray W. Johns, and **Caroline Dore.** Factor analysis of analogue scales measuring subjective feelings before and after sleep

Edward Stonehill, A. H. Crisp and **J. Koval.** The relationship of reported sleep characteristics to psychiatric diagnosis and mood

Richard Totman. Cognitive dissonance in the placebo treatment of insomnia—a pilot experiment

Volume 50 (1977) £16·00 ($39·50 *USA and Canada*). *Single parts* £5·00 ($11·75 *USA and Canada*), each

Enquiries about advertising should be sent direct to the Press.

CAMBRIDGE UNIVERSITY PRESS
BENTLEY HOUSE, 200 EUSTON ROAD, LONDON, NW1 2DB
32 EAST 57TH STREET, NEW YORK, N.Y. 10022

The British Journal of Psychology

Volume 67, Part 4, November, 1976

JOHN HINTON and BRIAN CRASKE. A relationship between Eysenck's P scale and change in muscle action potentials with attention to perceptual tasks
JEAN-YVES FRIGON. Extraversion, neuroticism and strength of the nervous system
MARTIN CRAWSHAW and BRIAN CRASKE. Oculomotor adaptation to prisms: complete transfer between eyes
J. ST. B. T. EVANS and P. C. WASON. Rationalization in a reasoning task
JOHN T. E. RICHARDSON. Procedures for investigating imagery and the distinction between primary and secondary memory
A. J. DURNDELL and N. E. WETHERICK. The relation of reported imagery to cognitive performance
BERNARD RIME and LEX McCUSKER. Visual behaviour in social interaction: the validity of eye-contact assessments under different conditions of observation
R. N. SYKES. A comment on Meddis' 'A simple two-group test for matched scores with unequal cell frequencies'
J. ST. B. T. EVANS. A critical note on Quinton and Fellows' observation of reasoning strategies
BRIAN J. FELLOWS. The role of introspection in problem-solving research: a reply to Evans
LAWRENCE PORTER. Rehearsal span versus remory span: what does this comparison tell us about verbal memory?
DAVID E. SOMEKH. The effect of embedded words in brief visual display
R. H. DAY and R. G. DICKINSON. The components of the Poggendorff illusion
J. SWELLER. The effect of task complexity and sequence on rule learning and problem solving
S. H. A. HENLEY. Responses to homophones as a function of cue words on the unattended channel
VICTOR M. AGRUSO, JR., and MARK D. RECKASE. The relative difficulty position function
DAVID S. BREE and GEORGE COPPENS. The difficulty of an implication task
MARIAN ANNETT. A co-ordination of hand preference and skill replicated
BOOK REVIEWS
OTHER PUBLICATIONS RECEIVED
EDITORIAL ACKNOWLEDGMENT

Volume 68 (1977) £20·00 ($49·50 USA and Canada). Single parts £6·00 ($14·00 USA and Canada) each
Enquiries about advertising in the journal should be sent direct to the Press.

CAMBRIDGE UNIVERSITY PRESS
Bentley House, 200 Euston Road, London, NW1 2DB
32 East 57th Street, New York, N.Y. 10022

THE BRITISH JOURNAL OF SOCIAL AND CLINICAL PSYCHOLOGY

Volume 15, Part 4, November, 1976

D. A. SHAPIRO. Reconstructing approximations to dialogue
CHRISTINE EISER and J. RICHARD EISER. Children's concepts of a fair exchange
CHRISTOPHER ORPEN. The expectancy beliefs, instrumentality beliefs and job reward valences of black and white workers in South Africa
L. B. HENDRY and D. J. E. THORNTON. Games theory, television and leisure: an adolescent study
BARRY McCARTHY and STEVEN W. DUCK. Friendship duration and responses to attitudinal agreement-disagreement
A. BEDFORD, G. A. FOULDS and B. F. SHEFFIELD. A new personal disturbance scale (DSSI/sAD)
DONALD K. FROMME, JOSEPH A. STOMMEL and RONALD D. DUVALL. Group modification of affective verbalizations: resistance to extinction and generalization effects
PHILIP LEY, PETER W. BRADSHAW, JOHN A. KINCEY and STEPHEN T. ATHERTON. Increasing patients' satisfaction with communications
P. D. SLADE. Towards a theory of auditory hallucinations: outline of an hypothetical four-factor model
DAVID A. T. SIDDLE, SARNOFF A. MEDNICK, A. R. NICOL and ROGER H. FOGGIT. Skin conductance recovery in antisocial adolescents
CLIVE READING and PETER D. MOHR. Biofeedback control of migraine: a pilot study
JESSE G. DELIA and BARBARA J. O'KEEFE. The interpersonal constructs of Machiavellians
DENISE C. R. BENEL and RUSSELL A. BENEL. A further note on sex differences on the semantic differential
ED. CAIRNS and BARBARA DURIEZ. The inuflence of speaker's accent on recall by catholic and protestant school children in Northern Ireland
G. F. REED. Indecisiveness in obsessional-compulsive disorder
BOOK REVIEWS
OTHER PUBLICATIONS RECEIVED

Volume 16 (1977) £19·00 ($47·50 *USA and Canada*). *Single parts* £6·00 ($14·00 *USA and Canada*) *each.*
Enquiries about advertising in the Journal should be sent direct to the Press

CAMBRIDGE UNIVERSITY PRESS
BENTLEY HOUSE, 200 EUSTON ROAD, LONDON, NW1 2DB
32 EAST 57th STREET, NEW YORK, N.Y. 10022

Biofeedback

The Mini-Relaxometer is a pocket-sized GSR feedback instrument providing the same function as larger Relaxometers through a single discreet earplug. It is useful wherever people need help in coping with stress, and particularly in anxiety management and in facilitating the transition between clinical and everyday-life situations.
Price £22.50 plus 8% VAT (cwo).

We have also widened the range of Cassettes which we distribute. Building on the foundation of CSW1, Relaxation, and CSW2, Stress Control, we now also offer CSW3, Thunder Stress, CSW4, Aircraft Flight Stress, CSW5, Agoraphobia, and CSW6, Interview Skills.

We also offer M1, Relax, an alternative approach to muscle relaxation with reference to migraine and effects of anxiety and excitement, by Jane Madders, Dip Phys Ed, MCSP. All cassettes cost £3,30 for one, £3 each for more, P&TP (cwo).

Details from
Aleph One Limited PO Box 72J Cambridge telephone 811679

PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

The British Journal of Psychology
The British Journal of Medical Psychology
Journal of Occupational Psychology
The British Journal of Social and Clinical Psychology
The British Journal of Mathematical and Statistical Psychology
The British Journal of Educational Psychology

The British Journal of Psychology, *The British Journal of Medical Psychology*, *Journal of Occupational Psychology* and *The British Journal of Social and Clinical Psychology* appear in quarterly parts. The 1977 subscription prices, payable in advance, are £20·00 (US $60·00 in the USA and Canada) for *The British Journal of Psychology*; £19·00 (US $57·00) for *The British Journal of Social and Clinical Psychology*; £16·00 (US $48·00) for *The British Journal of Medical Psychology* and £13·00 (US $39·00) for *Journal of Occupational Psychology*. Subscriptions to any of these journals may be sent to any bookseller or subscription agent, or to the Cambridge University Press, Bentley House, 200 Euston Road, London, NW1 2DB. American and Canadian enquiries should be addressed to the Cambridge University Press, American Branch, 32 East 57th Street, New York, N.Y. 10022.

The British Journal of Mathematical and Statistical Psychology is issued twice a year, in May and November. The subscription rate, payable in advance, is £15·00 per volume post free (US $34·00 in the USA). Subscriptions should be sent to the Society's agents, John Wright and Sons, Ltd., 42-44 Triangle West, Bristol, BS8 1EX, or to any bookseller.

The British Journal of Educational Psychology is issued three times a year, in February, June and November. The prepaid subscription is £8 a year, post free, or £3·00 each number. The prepaid subscription price in the USA is $20·00 post free. Subscriptions should be sent to the Scottish Academic Press, 33 Montgomery Street, Edinburgh, EH7 5JX.

The series of *Monograph Supplements* issued in connection with *The British Journal of Psychology* is now discontinued. Orders for back volumes may be sent to any bookseller, or the Cambridge University Press, Bentley House, 200 Euston Road, London, NW1, from whom details of these publications may be obtained.

MANUSCRIPTS

For *The British Journal of Psychology* should be sent to Professor A. D. B. Clarke, Department of Psychology, The University, Hull, HU6 7RX;

For *The British Journal of Medical Psychology* should be sent to Professor Arthur H. Crisp, Psychiatric Research Unit, Atkinson Morley's Hospital, Copse Hill, London, SW20 0NE;

For *Journal of Occupational Psychology* should be sent to Professor B. Shackel, Department of Human Sciences, University of Technology, Loughborough, Leicestershire.

For *The British Journal of Social and Clinical Psychology* should be sent to Dr. Halla Beloff (Social Psychology Editor), Department of Psychology, University of Edinburgh, 60 Pleasance, Edinburgh, EH8 9TJ; or to Dr. H. R. Beech (Clinical Psychology Editor), Department of Psychiatry, Withington Hospital, West Didsbury, Manchester, M20 8LR;

For *The British Journal of Mathematical and Statistical Psychology* should be sent to Professor P. Levy, Department of Psychology, University of Lancaster, Lancaster, LA1 4YF;

For *The British Journal of Educational Psychology* should be sent to Profesor N. J. Entwistle, Department of Educational Research, University of Lancaster, Lancaster.

Contributors are advised to consult *Suggestions to Authors* issued by the Standing Committee on Publications of the British Psychological Society and obtainable, price £0·50 (USA $2·00) post free, from Cambridge University Press.

All manuscripts should be submitted in duplicate.

OFF-PRINTS

Contributors on application receive fifty copies of their papers free. Additional copies may be obtained in units of fifty copies.

SUBSCRIPTIONS TO THE BRITISH PSYCHOLOGICAL SOCIETY

The subscription for Fellows, Associates and Graduate Members is £18·00 (£16·00 if paid by Banker's Order).

They are entitled to purchase any or all of the Society's Journals at the special rate of £4·00 per Journal per annum.

Information on membership and activities may be obtained from the Secretary, The British Psychological Society, St. Andrew's House, 48 Princess Road, Leicester, LE1 7DQ.

CONTENTS

SYMPOSIUM : READING ABILITIES AND DISABILITIES

H. FRANCIS : Children's Strategies in Learning to Read . . . 117

T. J. WHEELER, E. J. WATKINS and S. P. MCLAUGHLIN : Reading Retardation and Cross-Laterality in Relation to Short-Term Information Processing Tasks . . 126

R. J. RIDING and J. C. PUGH : Iconic Memory and Reading Performance in Nine-Year-Old Children 132

R. CONRAD : The Reading Ability of Deaf School-Leavers . . . 138

ARTICLES :

J. M. ENGELHARDT : Analysis of Children's Computational Errors: A Qualitative Approach 149

J. ZAMMARELLI and N. BOLTON : The Effects of Play on Mathematical Concept Formation . 155

M. GHODSIAN and M. CALNAN : A Comparative Longitudinal Analysis of Special Education Groups 162

C. GOLOMB : Symbolic Play: The Role of Substitutions in Pretence and Puzzle Games 175

G. M. SEDDON : The Effects of Chronological Age on the Relationship of Academic Achievement with Extraversion and Neuroticism: a Follow-up Study 187

RESEARCH NOTES :

W. S. ANTHONY : The Development of Extraversion and Ability: An Analysis of Rushton's Longitudinal Data . . . 193

E. CAIRNS : The Reliability of the Matching Familiar Figures Test . 197

E. SCOTT : Treatment of Encopresis in a Classroom Setting : A Case Study 199

A. J. MASSEY : Candidate Fatigue and Performance on G.C.E. Objective Tests 203

C. V. NEWMAN, R. COCHRANE and D. E. BLACKMAN : The Role of Interviews in Choice of University by Psychology Undergraduates 209

BOOK REVIEWS 215

PUBLICATIONS RECEIVED 231

UK ISSN 0007-0998

THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

NOVEMBER 1977

VOL. 47, PART 3

THE
BRITISH JOURNAL
OF
EDUCATIONAL PSYCHOLOGY

Edited by
N. J. ENTWISTLE

Assistant Editors :
S. N. BENNETT
G. BROWN

Sub-Editor :
S. M. CUNNINGHAM

Advised by
L. B. BIRCH
E. A. PEEL
P. E. VERNON

Vol. 47

Published by
SCOTTISH ACADEMIC PRESS, 33 MONTGOMERY STREET, EDINBURGH, EH7 5JX,
for the British Journal of Educational Psychology, Ltd.

Registered Offices: 8-24 Sampson Road North, Birmingham, B11 1BD. This address should not be used for letters or books intended for the Editor, Authors, or the Publishers.

On behalf of
The British Psychological Society.

All rights reserved.

© 1977. The British Journal of Educational Psychology.

PRINTED BY BIRMINGHAM PRINTERS LTD., 8-24 SAMPSON ROAD NORTH, BIRMINGHAM.

Br. J. educ. Psychol., 47, 233-243, 1977

SYMPOSIUM: LEARNING PROCESSES AND STRATEGIES — III

ON QUALITATIVE DIFFERENCES IN LEARNING: III — STUDY SKILL AND LEARNING

By L. SVENSSON

(Institute of Education, University of Göteborg, Sweden)

SUMMARY. Detailed information was collected by interview from a sample of 30 first-year education students, who took part in a learning experiment and described their approaches to their normal studies. It was found that the distinction between holistic and atomistic cognitive approaches was consistent over different occasions when students were asked to read and to recall their knowledge of the two texts in the experiment. This concept of cognitive approach also had a functional relationship with academic attainment, helping to explain not only examination success, but also other aspects of the students' approaches to studying.

INTRODUCTION

Previous articles in this symposium (Marton and Säljö, 1976a, 1976b) have examined the relationship between process and outcome in reading an academic text, and the effect of different types of questions on the learner's characteristic approach to learning. The distinction between deep-level processing and surface-level processing emerged from these analyses of students' qualitative differences in their conception of the author's intended meaning. One of the important advantages of this research strategy is that it becomes possible to examine the processes of learning in sufficient detail to be able to *explain,* rather than simply describe, the measured outcomes. The present article capitalises on this advantage in examining the concept of *study skill.*

First it is important to clarify how the research strategy used in these studies of students' learning differs from traditional approaches. It is argued that previous research in the area has failed to establish *functional* relationships between study activities and learning outcomes. There have been many studies which have demonstrated *correlations* between various aspects of study methods, habits or attitudes and academic performance (for example, Biggs 1970; Brown and Holtzman 1955; Entwistle *et al* 1971; Goldman and Hudson 1973). Such correlational analyses average out interesting differences between studies to present a misleading description of a general relationship, providing only weak evidence from which to understand how study activities relate to academic attainment. Nor do more recent approaches using a.t.i. (Biggs 1976) or cluster analysis (Entwistle and Wilson 1977), *in themselves,* provide sufficient information.

The main weakness in all these approaches is that the variables are theoretical constructions imposed on the phenomenon. If our understanding is to be close to that phenomenon, then the concepts used to describe and to understand it should emerge from analysing instances of the concept observed in different people and in different situations or contexts. It is considered essential to recognise that a student's *perception* of a situation is as important as the situation itself in understanding a relationship between concepts. It is also necessary to examine relationships between instances of concepts described by the same individual on different occasions, as well as the patterns of similarities and differences found between groups of individuals categorised in the same way. Only when data are collected and analysed in this way can a clear indication of the relative importance of different functional relationships be fully understood.

The intention in the present study was thus to collect instances of study skill in different situations, and to relate study activity to levels of understanding and academic performance. The situations chosen for this purpose were first a learning experiment designed to demonstrate qualitative differences in learning from an academic text, and second the students' approaches to normal studying in relation to their success in examinations. An additional aspect of this study was a reanalysis of the data described by Marton and Säljö (1976a) which led to the concepts of deep- and surface-processing as explanations of qualitative differences in what students came to understand from reading an academic text. As the identification of categories and concepts involves subjective interpretations of the data, it seemed essential to check the inter-researcher agreement, not in the relatively trivial sense of agreed assignment to predetermined categories, but in terms of the extent to which two investigators, with theoretical perspectives in common, would select similar, or different, aspects of the students' responses as being most significant.

METHOD

Sample

Data were collected over a period of one term from 30 volunteer first-year education students who were following five courses. The students ranged in age from 19 to 23 years and 25 of the students were women. This sample is identical with that reported by Marton and Säljö (1976a), but additional data were collected.

Procedure

Data consisted in responses to questions about what had been learned, or could subsequently be recalled, from learning experiments, and also information about the students' normal approach to studying and their examination performance. Using the procedure described by Marton and Säljö (1976a) students were individually asked to read, without time limit, an academic text and were subsequently questioned to reveal their understanding of the author's meaning, and their grasp of subsidiary aspects of the learning material. Students were also asked to provide introspective reports about how they had tackled the task of reading the article. After an interval of five weeks the students were contacted again and asked to report what they remembered of the text. At this second session new information was also collected. After recalling what they remembered of the first text, students were asked to read a second passage, which was, in fact, an extended version of the first one containing additional information and an extension of the argument. Finally, after another interval of five weeks, students were asked to recount what they remembered of the second article. Each of the three test sessions ended with questions about the students' normal study activities.

The intention in collecting these data was to delimit the concept of *study skill* first in terms of a controlled, specific learning situation (the experiment) and then in relation to the total pattern of study shown in normal academic work. In the experiment the criterion was one of understanding, described in terms of the level of understanding reached. This understanding was related to the student's cognitive approach which Marton (1975) had described previously in terms of differing levels of processing. The criterion in normal studies was examination performance which could, in turn, be related again to cognitive approach but also to other aspects of the student's pattern of study. The latter consists of attendance record, time spent on independent study, amount of course literature read, distribution of work over the term, and revision. Fuller details of the procedure have been described elsewhere (Svensson, 1976). The research design is shown, in outline, in Figure 1.

Session 1		Session 2		Session 3
Reading Text 1 Recalling Text 1 Introspective reports Specific questions on Text 1 Questions about normal studies	Interval of five weeks	Recalling Text 1 Introspective reports on remembering Reading Text 2 Recalling, introspective reports and specific questions on Text 2 Questions about normal studies	Interval of five weeks	Recalling Text 2 Introspective reports on remembering Specific questions on Text 2 Questions about normal studies

Figure 1. Outline of the research design.

Analysis of data

In accord with the position taken in the introduction, no quantitative or statistical analyses of the data were undertaken. The aim was to *delimit* the concept of study skill in relation to its specific context. Thus the data were analysed in terms of the functional relationship between outcome (criterion) and learning or studying activities identified *within* each individual. Relationships between variables based on mean scores from groups of individuals were not examined. Nevertheless it is possible, by induction from the functional relationship identified in the data from each individual, to build up a theoretical understanding of a more general meaning of study skill, while retaining the view that a full understanding can only be achieved in relation to the study activities and learning outcomes of individual students.

Three main analyses were undertaken. The first re-examined the data from reading Text I in terms of qualitative differences in outcome and process of learning. The second analysis was concerned with the effects of recalling Text I on the reading and recalling of Text II. Finally a student's characteristic approach revealed in the learning experiment was related to his normal pattern of studying and his examination results. As much of the analysis of the learning experiment depends on the qualitatively different meanings students derived from reading the texts, the content of the passages must first be described in some detail.

Texts

Both texts were written by one of our colleagues (Dahllöf, 1968, 1969), the second being a fuller exposition of the first. The first text has already been described but, for convenience, the summary is repeated here.

> "The newspaper article was 1,400 words long and included three tables. The article was mainly a critique of the approaching curriculum reform in

> the Swedish universities (UKAS), which aimed at bringing studies more into line with those at the polytechnic institutes through the introduction of set combinations of subjects and stricter regulations as regards duration of studies (termination in the case of unsatisfactory examinations results). The reason for the reform, as explained by the authorities, was that the examination pass rate at the universities was considerably lower than that achieved at the polytechnic institutes.
>
> The author of the article had, after examining the underlying statistics, divided university students into sub-categories and was thereby able to show that, even though the pass rate was very low for certain categories of students, for other categories it was as high as, if not higher than, that achieved by technical students. The author argued that the blanket approach of the university reform, which would affect all equally, was misguided. If the pass rate was to be raised (and this was not considered self-evident by the author) selective measures should be taken by concentrating on those groups that did have a low pass rate'' (Marton and Säljö, 1976a, p8).

The second text was approximately three times longer than the first. With minor amendments, all the main points and structure of the first text were retained, but with an additional, substantial section on how long students had been studying. This section clarified the explanations of the differences in examination pass rates, which provided the basis for the main conclusions.

Both texts were carefully chosen to ensure that the students would possess the necessary prerequisites for understanding, in terms of linguistic structures, terms and concepts used. Thus 'these aspects of learning as a skill were controlled by careful choice of the level and content of the texts in relation to the students' previous experience.

RESULTS

Qualitative differences in outcome

Immediately after reading the text students were asked to relate everything they remembered about it. As before (Marton and Säljö, 1976a) the first step was to examine different conceptions of the intentional content of the article, as reported by the students. But, in this case, the interpretation of the data was carried out, independently, by the present author.

Marton and Säljö had reported that the main qualitative differences in outcome could be described by four distinctive ways of describing the message contained in Dahllöf's newspaper article about curriculum reform in Swedish universities. The different levels represented decreasing accuracy in recounting the main argument presented.

Level A	Selective measures are needed—only for those groups of students who did not fulfil the requirements.
Level B	Differential measures are needed—to allow for differences between various (unspecified) groups.
Level C	Measures are needed—without reference to their being selective or differential.
Level D	There are differences between groups of students attending university—without any indication of the meaning or implications of those differences.

The independent analysis of students' levels of understanding of the first text confirmed the above categorisation, and there was close agreement in assigning students to categories in terms of their level of understanding. There was exact agreement in 26/30 cases for immediate recall and in 24 cases for delayed recall. Where there was disagreement, only in one case (for delayed recall) was there more than one category of difference.

After the students had read the second text, a new category appeared, which was based on a rather different aspect of the author's argument. In terms of level of understanding, it falls between the previous categories B and C.

Level BC The efficiency of the universities is not so bad as the authorities believe.

Furthermore, after reading the second text, it was necessary to sub-divide levels B and C according to whether the students' comments related to the organisation of university studies, or to the way investigations concerning the efficiency of university education should be carried out.

A comparison of the levels of outcome showed a clear structural difference between levels A, B and BC on the one hand, and C and D on the other. Only in the first three categories did the students make use of the author's linkage between the facts presented and his conclusion. In categories C and D this relation between facts and conclusion is not present in the students' recall. If the data are reorganised into a simple dichotomy—conclusion/no conclusion—for Text I there is agreement in the two independent categorisations in 29/30 cases.

Cognitive activity

The introspective reports of the students allow learning processes, as well as learning outcomes, to be examined. The result of previous analyses was to describe differences in cognitive activity and attitude in terms of surface- and deep-level processing.

> "In the case of *surface-level processing* the student directs his attention towards learning the text itself *(the sign),* ie, he has a 'reproductive' conception of learning which means that he is more or less forced to keep to a rote-learning strategy. In the case of *deep-level processing,* on the other hand, the student is directed towards the intentional content of the learning material *(what is signified),* ie, he is directed towards comprehending what the author wants to say about, for instance, a certain scientific problem or principle" (Marton and Säljö, 1976a, pp7-8).

The present author identified similar differences between students' approaches to the texts, but the data were examined from the point of view of shared structures between knowledge and skill. Previously, emphasis has been placed on one or other of these aspects of learning—traditionally on the knowledge acquired, but more recently on the importance of 'learning to learn' as a skill, and on the necessary connection between knowledge and this skill.

> "The acquisition of knowledge as the primary goal of education must be reconsidered. The analysis we have developed points to the centrality of skill acquisition. To neglect skill is to forget about how an individual acquires and uses knowledge. Formal education has always been concerned with knowledge because it is tied directly to the content of any medium; it has tended to overlook skills because they are implicit in the very use of the media of expression, communication and instruction. How to adjust the forms of instruction to the joint development of knowledge and skill—that is the issue for further debate." (Bruner and Olsen, 1974, p149).

The stand-point taken in analysing the present data emphasises, even more strongly, the unity of knowledge and cognitive skill. It is argued that the structural properties of an individual's conception of various phenomena are the most important aspects of his knowledge, because they also constitute fundamental characteristics which determine how he acquires and uses knowledge.

When the students' reports of how they tackled the reading of the texts were related individually to what the same students came to understand of the meaning of those texts at each of the three test sessions, the processes relating to the 'conclusion/no conclusion' outcome became clear. Within the knowledge/skill frame-work of interpretation, the differences in approach to the first text were described as 'atomistic' contrasted with 'holistic' cognitive approaches.

The *atomistic* approach was indicated when students described their activities as involving: focusing on specific comparisons, focusing on the parts of the text in sequence (rather than on the more important parts), memorising details and direct information indicating a lack of orientation towards the message as a whole. In contrast the *holistic* approach was characterised by students' attempts: to understand the overall meaning of the passage, to search for the author's intention, to relate the message to a wider context and/or to identify the main parts of the author's argument and supporting facts.

These two types of activity were also apparent in students' reports on how they tried to remember the texts. 'Atomists' relied on remembering the introductory sentences, visualising the tables, parts of the text or the outline structure of the text, and/or a general orientation to details. 'Holists' mentioned their attempts to remember the main message, what the author had been trying to say, the basic steps in the argument, and the message in a wider context.

These descriptions parallel the distinctions between deep- and surface-processing made by Marton and Säljö, but the terms used here are chosen to avoid a distinction between outcome and process. Knowledge and skill, in the present terminology, are seen as unified. However, when the two terminologies were applied separately to the data, there was close agreement when students were classified as deep/holist, surface/atomist or unclassifiable. In 25/30 cases the same category was chosen, and in no case was a student placed in 'opposite' categories.

The descriptions of holist and atomist approaches should, of necessity, relate closely to the 'conclusion/no-conclusion' outcomes identified earlier. It has been suggested that there should be a close agreement between knowledge and skill. Table 1 shows the degree of similarity in categorisation on the various occasions that students were asked to recall what they remembered of the text.

TABLE 1

RELATIONSHIP BETWEEN COGNITIVE APPROACH DURING LEARNING AND CATEGORY OF OUTCOME FOR IMMEDIATE AND DELAYED RECALL OF EACH TEXT

	Text 1* Recall				Text 2 Recall			
Cognitive Approach	Immediate		Delayed		Immediate		Delayed	
	Concl	No concl	Concl	No concl	Concl	No concl	Concl	No concl
Holistic	14	1	8	0	16	1	14	0
Atomistic	1	13	7	14	3	10	5	11

*One student has been omitted from this analysis due to missing information caused by a recording fault, which prevented categorisation on cognitive approach, but not outcome.

In Table 1 it is particularly striking that for each of the two texts there was only one student, categorised as adopting a holistic approach, who failed to demonstrate in the knowledge structure of his learning outcome a grasp of the facts/conclusion relationship in the author's message. After five weeks, particularly in relation to Text 1, fewer holistic approaches could be identified, with only 8/15 students who recognised the facts/conclusion relationship showing clear indications of a holistic approach. It is also apparent that substantially more students could be recognised as retaining a holistic approach on delayed recall after reading the longer, second text. The relatively small number of holists identified at delayed recall of the first text could thus be a result of the shorter text used in that experiment.

Table 1 is taken as providing substantial evidence for the conclusion that cognitive approach, as consciously recognised and reported by the students, bears a close relationship to the knowledge understood. Holists, with only one exception in each experiment, reached a deep level of understanding of the author's intended meaning. The tendency for a substantial number of students classified as adopting atomistic approaches on delayed recall to reach fairly deep levels of understanding may reflect a weakness in clearly identifying holists on these occasions, rather than a weakening in the association between knowledge and the cognitive approach of these students. To examine this possibility more closely, it is necessary to move away from the group tendencies shown in Table 1 and examine the intra-individual consistencies in the categories used to describe each individual on different occasions.

Outcome of learning on different occasions

Levels of understanding were identified on four occasions (immediate and delayed recall for each text) and on each occasion the 'conclusion/no-conclusion' (C/NC) dichotomy was recorded. Table 2 presents the intra-individual variations in these categories across occasions.

TABLE 2

INTRA-INDIVIDUAL PATTERNS IN THE CLASSIFICATION OF LEVELS OF OUTCOME ON FOUR OCCASIONS

Group Number	Classification Pattern				Number of individuals so classified
	Text 1 Immed	Text 1 Delay	Text 2 Immed	Text 2 Delay	
1	NC	NC	NC	NC	10
2	C	C	C	C	7
3	C	NC	C	C	7
4	C	NC	C	NC	1
5	C	NC	NC	NC	1
6	NC	NC	C	NC	3
7	NC	NC	C	C	1

Note: C represents a recognition of the facts/conclusion relationship in the text.
NC indicates a failure to show an awareness of this relationship.

Table 2 shows that 25/30 students were placed in the same category of outcome on immediate recall, and 17/30 on all occasions. Analyses of the extent to which students showed the same category of level of understanding when asked to recall specific parts of the text on the various occasions (Svensson, 1976), together with the pattern of results shown in Table 2, indicate that consistency in cognitive approach, rather than memory of the meaning of factual information, provides the best explanation of the findings. Of the 16 students who recognised the facts/conclusion relationship in Text 1, all but one also showed this understanding on immediate recall of Text 2 and there were special reasons why this one student did not learn effectively on the second occasion. Of the 14 students who did not recognise the facts/conclusion structure in Text 1, only 4 acquired it after reading Text 2 (3 at the marginal BC level), and only one retained that understanding over the final five-week interval.

If students are consistent in their cognitive approach within the learning experiment, to what extent do they use similar processes in their normal study activities and how are these different approaches related to examination performance?

Cognitive approach, study technique and academic success

In the interviews students were asked questions about their typical approaches to studying. Again in the replies were clear indications of differences in cognitive approach, but the indications were, necessarily, less precise than those obtained in the specific situation of the learning experiment. Students adopting a holistic approach to their normal studies related new material to their own knowledge and experience, stressing the importance of reorganising new information in terms of existing knowledge structures. For example, one student talked about the problems of reacting appropriately to examination questions and thereby showed a holistic orientation to studying.

> "One builds a sort of construction, a frame you could say, figuratively speaking. A book—one has a frame for it and to that certain associations; and then come the questions that do not—absolutely do not fit in. The whole thing collapses, and then one has to begin again and build a new structure that will stand up to those questions, and there you have it—it has changed."

The atomists were more likely to stress the importance of over-learning and memorising in preparation for examinations.

> "*E:* Do you try to memorise certain things more than others?
> *S:* Yes.
> *E:* What sort of things?
> *S:* Important things, like these names and these numbers, I think it is essential to learn them . . . "

When the students were categorised into holists and atomists from their comments during the three interview sessions, there was a clear indication of a functional relationship between cognitive approach, in both the experiment and normal studies, and examination performance.

TABLE 3
RELATIONSHIP BETWEEN COGNITIVE APPROACH AND EXAMINATION PERFORMANCE

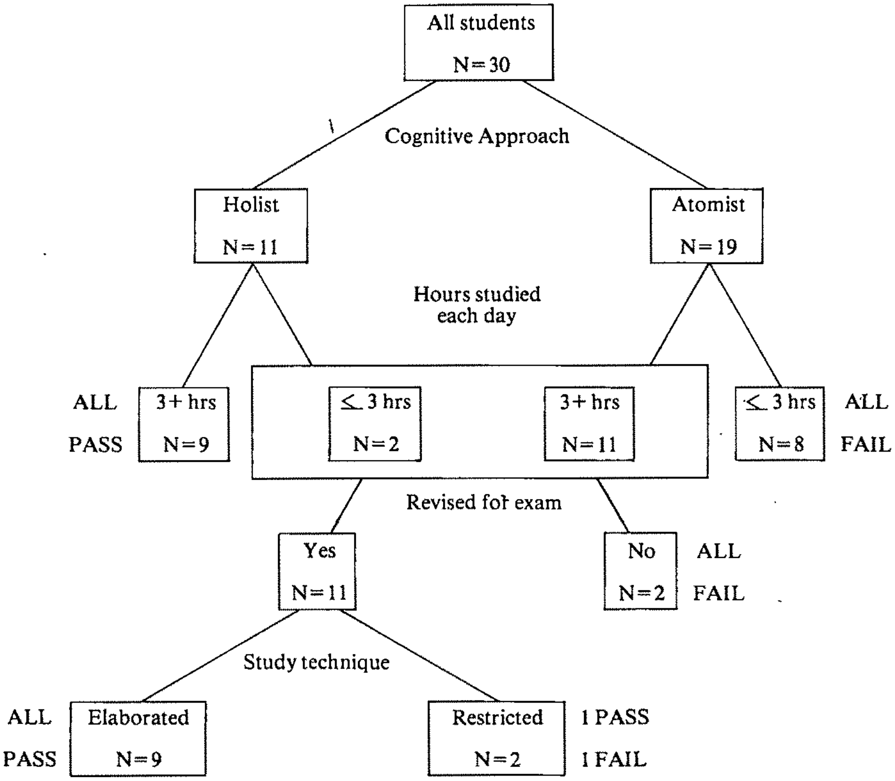
Cognitive Approach		Examination Performance		
Experiment	Normal studies	Passed All	Some Failure	Total
Atomistic	Atomistic	3	10	13
Holistic	Holistic	9	1	10
Holistic	Atomistic	4	2	6
Atomistic	Holistic	1	0	1

Table 3 shows, firstly, that there was substantial agreement between categorisations of cognitive approach in the experiment and in normal studies. 23/30 students were placed in the same category, even though there was some uncertainty in identifying atomistic approaches to normal studies. There were marked differences in the examination performance of students categorised as holists and atomists in the experiment. 13/16 holists passed all examinations, compared with 4/14 atomists. Among those students who were identified as holists in the experiment and in their normal studies, only 1/10 showed any element of failure in the courses taken in the first term. It appears, therefore, that difference in cognitive approach provides a good, and functional, explanation of academic performance in examinations, although it is not, in itself, a full explanation.

The advantage of collecting detailed information about a group of students is that it is also possible to explore functional relationships between cognitive approach and other aspects of studying. Students gave information about study technique, the time spent on average each day on private study, and how much revision they had done prior to each examination. (Additional aspects were also explored—see Svensson, 1976.) Analysis of instances of study technique suggested that the main distinction was between students who adopted an *elaborated,* as opposed to a *restricted,* technique. An elaborated technique was indicated by making summaries, underlining important parts of books, and more use of lecture notes. Bringing together all the information on normal studying allows a much more complete explanation of academic performance to be presented. Figure 2 indicates the very high level of successful differentiation between students who passed or failed one of the education examinations (29/30 correctly identified) using four dichotomised concepts..

FIGURE 2

DIFFERENTIATION BETWEEN STUDENTS WHO PASSED AND FAILED AN EXAMINATION IN EDUCATION IN RELATION TO INFORMATION ABOUT THEIR APPROACHES TO STUDYING



The advantage of examining instances of concepts within, as well as between, individuals also enables the more important functional relationships to be identified. For example, how should the association between atomistic approach and number of hours studied be interpreted? It was clear from the interviews that adopting an atomistic approach created problems for students. The type of learning demanded in higher education becomes exceedingly difficult to tackle in an atomistic manner. Memorisation of the many books, articles and lecture notes is an impossible, as well as unnecessary, task. Some students adopting an atomistic approach are sufficiently motivated to accept the long hours of tedious work necessary to learn this way. For others this approach makes the learning boring and irrelevant, and hence the pattern of declining effort and increased incidence of examination failure is understandable. This is another reason why cognitive approach is seen as having such an important functional relationship with academic success.

DISCUSSION

It was agreed earlier that a clear understanding of the importance of study skills depended on recognising the importance of context, and academic content, in defining those skills. It was also suggested that an examination of functional relationships could not be achieved by traditional approaches to this problem. The use of extensive interviews about students' approaches to reading a specific text, and to their normal studying, enables a delimitation of study skills to emerge in two distinct contexts in relation either to specific knowledge content or to specific academic courses. Examining the intra-individual patterns of cognitive approaches and learning outcomes (Tables 2 and 3), as well as inter-individual relationships between cognitive approaches and examination results, allows a clearer explanation of the functional relationships between study skills and academic performance to emerge.

It seems clear from the findings that there is a sense in which there is an essential unity between skill and knowledge in learning. This unity depends on a common structure of cognitive activity and learning outcome. If a deep understanding of the author's meaning is to be reached, a holistic approach is essential. In the learning experiment, the students who recognised the fact/conclusion structure must have looked out for such a relationship, since it cannot have been known in advance. The structure is implicit, rather than explicit, in the text and cannot be recognised by students adopting an atomistic approach who conceptualise the message as no more than a series of associated parts. A holistic learning process is thus a necessary prerequisite for the acquisition of a deep level of understanding.

This unity does not, however, necessitate identical approaches being used under all conditions, or with all materials. The emphasis on context in a phenomenological sense implies that there will be intra-individual differences in cognitive approach in different situations, depending on how the student conceptualises what is required of him. The students interviewed had probably all used conclusion structures in some, though not necessarily academic, contexts. It is probably an important characteristic of cognitive development that a holistic approach will be extended to more and more complex units of information. In the present study only 16/30 students had utilised the holistic approach in the learning experiment and only 10 of those 16 had clearly extended this approach into their normal studying.

The inability of many students to utilise a holistic approach in reading academic articles and in normal studying is particularly worrying in two ways. First the unity of cognitive approach and knowledge structure indicates that the adoption of an atomistic approach effectively prevents a deep level of understanding being reached, and second, the close association between cognitive approach and examination performance indicates than an atomistic approach is likely to lead, for a large majority of students, to at least partial failure in their courses. In other analyses (Svensson, 1976) the cumulative disadvantages of an atomistic approach were demonstrated over a full academic year for students taking social sciences.

The present study allows no more than speculation about the origins of the different cognitive approaches identified and their educational implications. It seems probable that the emphasis on knowledge of facts in secondary school examinations is at least partly responsible for students' approaches to learning in higher education. It may be difficult for them to realise that there are, or at least should be, different demands made in higher education. The development of personal knowledge and critical thinking are commonly thought to be an essential outcome of higher education. Yet only by adopting a holistic approach is the student led to relate and reflect on items of knowledge acquired. The atomistic approach leaves the student uncritically accepting and memorising separate parts of what they find in books or in lectures. It seems essential to help students who adopt atomistic approaches to recognise the importance of interacting actively and critically with the information and ideas presented in the course of higher education. Otherwise it is difficult to see what lasting effects the experience of higher education could have.

ACKNOWLEDGMENTS.—The research reported here has been financed by the R and D Unit of the National Swedish Board of Universities and Colleges. I owe a debt of gratitude to Noel Entwistle for creating Figure 2 and for advising me in the preparation of the article.

REFERENCES

BIGGS, J. B. (1970). Faculty patterns in study behaviour. *Aust. J. Psychol.*, 22, 161-174.

BIGGS, J. B. (1976). Dimensions of study behaviour: Another look at ATI. *Br. J. educ. Psychol.*, 46, 68-80.

BROWN, W. F., and HOLTZMAN, W. H. (1955). A study — attitudes questionnaire for predicting academic success. *J. educ. Psychol.*, 46, 75-84.

BRUNER, J. S., and OLSON, D. R. (1974). Learning through experience and learning through media. In OLSON, D. R. (Ed.), *Media and Symbols: The Forms of Expression, Communication and Education.* Chicago: The University of Chicago Press.

DAHLLÖF, U. (1968). Examensfrekvens, studentkategori och lärosäte. En omprövning av frågan om de fria fakulteternas effektivitet. (Examination pass rate, student category and university. A reconsideration of the question about the efficiency of the open faculties.) *Rapporter från* Pedagogiska institutionen, Göteborgs universitet, nr 31.

DAHLLÖF, U. (1969). Examensfrekvens, studentkategori och lärosäte. (Examination pass rate, student category and university.) *Nordiskt Forum,* 1969:4.

ENTWISTLE, N. J., NISBET, J., ENTWISTLE, D., and COWELL, M. D. (1971). The academic performance of students. 1 — Prediction from scales of motivation and study methods. *Br. J. educ. Psychol.*, 41, 258-267.

ENTWISTLE, N. J., and WILSON, J. (1977). *Degrees of Excellence: the Academic Achievement Game.* London: Hodder and Stoughton.

GOLDMAN, R. D., and HUDSON, D. J. (1973). A multivariate analysis of academic abilities and strategies for successful and unsuccessful college students in different major fields. *J. educ. Psychol.*, 65, 364-370.

MARTON, F. (1975). On non-verbatim learning: 1. Level of processing and level of outcome. *Scand. J. Psychol.*, 16, 273-279.

MARTON, F., and SÄLJÖ, R. (1976a). On qualitative differences in learning: 1. Outcome and process. *Br. J. educ. Psychol.*, 46, 4-11.

MARTON, F., and SÄLJÖ, R. (1976b). On qualitative differences in learning: II. Outcome as a function of the learner's conception of the task. *Br. J. educ. Psychol.*, 46, 115-127.

SVENSSON, L. (1976). *Study Skill and Learning.* Göteborg: Acta Universitatis Gothoburgensis.

(Manuscript received November 20, 1976)

Br. J. educ. Psychol., 47, 244-257, 1977

ON QUALITATIVE DIFFERENCES IN LEARNING: IV—EFFECTS OF INTRINSIC MOTIVATION AND EXTRINSIC TEST ANXIETY ON PROCESS AND OUTCOME

By A. FRANSSON

(Institute of Education, University of Göteborg, Sweden)

SUMMARY. Eighty-one students were asked to read an article under different conditions of extrinsic and intrinsic motivation. The conditions were varied by choosing the sample so as to make the article relevant or irrelevant to contrasting subgroups, thus attempting to control the level of intrinsic motivation. The test conditions were varied to increase ego-involvement and threat to self-esteem in one situation, while providing a situation for other subgroups which was supportive and non-demanding. The qualitative differences in learning process and outcomes and the quantitative differences in recall of factual knowledge were investigated in relation to the various experimental subgroups created. Self-reports on trait and state test-anxiety, and the extent to which the students had actually experienced the types of motivation intended to be produced by experimental manipulation, were also investigated in relation to qualitative and quantitative differences in learning. Lack of interest in the text, efforts to adapt to expected test demands, and high test anxiety, were all found to increase the tendency towards surface-processing and ineffective, reproductive attempts at recall. However, an adaptive approach allied to strong interest and low anxiety produced a high proportion of deep-level approaches with good factual recall.

INTRODUCTION

In the preceding articles in this symposium results from the TIPS project (Study skill and learning) at the University of Göteborg have been presented and discussed (Marton and Säljö, 1976a, 1976b; Svensson, 1977).

Important points of departure for this project have been the convictions that "Both for instructional purposes and for the understanding of 'what it takes to learn' a description of *what* the students learn is preferable to the description of *how much* they learn" (Marton and Säljö, 1976a) and that in order to be able to describe what is learned "one must take into account *the content* of the learning or the discourse".

The main research method used was experiments where the process and outcome of learning of small numbers of subjects was intensively studied. Subjects were asked to read one or more passages of academic prose and were subsequently asked questions about the passage and about how they approached the text.

In each study of the experimental series a number of hierarchically ordered categories containing basically different conceptions of the meaning of the passages read could be identified. These categories were called *levels of outcome.* (Marton, 1975a; Säljö, 1975; Dahlgren, 1975; Svensson, 1976). Two levels of processing were also repeatedly identified. Marton and Säljö (1976a) describe these levels as *surface-level* and *deep-level processing* while Svensson (1977), from a slightly different point of view, prefers to talk about an *atomistic* compared with a *holistic* approach. A close relation was found between level of processing and level of outcome.

Efforts to induce deep-level processing by content neutral questions (Marton, 1974; 1975b), content-oriented guidance inserted in the text

(Dahlgren, 1975) and different types of test questions (Säljö, 1975) were not altogether successful. The result of the experimenters' efforts to induce deep-level processing in many cases turned out to be a narrow, technified learning process.

After an analysis of the perplexing results of Dahlgren and Säljö and of studies of the reading process, Edfeldt (1976) concluded that every form of disruption of the continuous contact between the reader and the content of the text has negative effects on the outcome of learning. There is no short-cut to deep-level processing.

In spite of the failure to induce deep-level processing, these experiments are very interesting. They clearly demonstrate how the subjects adapt their learning approach to their beliefs of what is required of them. The subjects prepare for what they expect to be the performance requirements.

Effort to satisfy external requirements (extrinsic reward) is not the only kind of motivation for learning, however. Spontaneous interest or intrinsic motivation is another kind. The purpose of the experiment reported in this article was to study simultaneous effects of these two kinds of motivation on learning.

Teachers and pupils as a rule believe motivation to be one of the most important factors affecting learning. Experimental studies on the relation between learning and motivation, however, have not given much support to this belief. The behaviouristic research tradition with its emphasis on quantitative aspects of behaviour has focused on the intensity dimension of motivation and produced results indicating both positive and negative correlations between intensity of motivation and learning. The curvilinear relationship summarising the well-known Yerkes-Dodson law is often relied upon to provide a comfortable explanation. But Biggs (1971) argues that the Yerkes-Dodson law only describes the relation between *extrinsic* motivation and learning. Saltz (1971) shows that the curvilinear relation might be regarded as a combined effect of at least two different strategies for defending against the threat associated with extrinsic motivation.

Several researchers, from Dewey (1913) on, have underlined the importance of making a distinction between intrinsic and extrinsic motivation. Bruner (1961) and White (1959) contrast drive-reduction with 'competence' motivation. DeCharms (1961) and Lefcourt (1976) distinguish between internal and external locus of control.

It is argued by these researchers, and many others, that intrinsic motivation causes distinctly different kinds of learning compared with extrinsic motivation, and that there is no 'optimal level' of motivation in the relation between intrinsic motivation and learning. In spite of these very important theoretical considerations there seem to be few, if any, experimental studies simultaneously demonstrating effects of intrinsic and extrinsic motivation. The reason for this is not hard to understand. Parknäs (1976) summarised the main reason. "Intrinsic motivation is not to be created. It is to be found." Intrinsic motivation is not open to experimental manipulation. The experimenter has to find subjects with real differences in intrinsic motivation for studying the learning material he has prepared for his experiment.

The problem of operationalising contrasting types of motivation has been tackled in the present study, beginning with formal definitions to highlight the distinctions.

Intrinsic motivation for learning is a state where the relevance for the learner of the content of the learning material is the main reason for learning.

Extrinsic motivation for learning is a state where the reasons for the learning effort have nothing to do with the content of the learning material. A good learning performance serves merely as a means for achieving some desired end result.

The following hypotheses provide the starting point of the study:

(1) When learning a text is made instrumental to reach important goals or to avoid aversive consequences with no relation to the content of the text, the student will try to memorise the text in order to be able to repeat the whole passage. He will make his greatest efforts on the points that he thinks that the examiner will question him about. If he has got no hints on probable questions he will follow the outline of the text. The greater his external motivation the more effort he will put in his reading.

(2) In contrast, when learning a text is not instrumental to reach goals or avoid aversive consequences, the student will approach the text in a personal way. He will not try to memorise the text but rather use it for his personal needs, which means that he will read selectively and focus on the meaning of those aspects of the text that he finds interesting.

(3) If most of the content is irrelevant to his interests and needs his intrinsic motivation for reading the text will be very low and he will have little to say about the text than just that it was not interesting. If he is asked to answer test questions on the contents of the text he will perform very badly. If on the other hand the content of the text meets important needs for information, his intrinsic motivation will make him absorbed in the text and he will learn effectively even without a conscious intention to do so.

(4) When learning achievement affects the self-esteem and social value of the student, level of test anxiety will be an important factor determining the quantitative outcome of learning. Students low in test anxiety will perform better under external motivation than those with higher levels of anxiety (Costello, 1964; Lekarczyk and Hill, 1969; Sarason, 1975).

(5) Students with strong intrinsic motivation for reading a text will be put into a conflict between different learning strategies if learning the text is made instrumental to reach some goal irrelevant of the content of the text. This conflict will cause them to learn less than extrinsically motivated subjects with low intrinsic motivation.

METHOD

Choice of learning material and choice of subjects

Level of extrinsic motivation is open to experimental manipulation. Level of intrinsic motivation is not. In order to be able to investigate effects of level of intrinsic motivation on learning the experimenter has to choose learning material that some of his subjects find relevant and other subjects find irrelevant to their personal interests and needs. Choice of learning material and choice of subjects are thus closely related. In this experiment a solution to this problem was sought in the following way.

A text containing a description of the examination system at the Institute of Education, University of Göteborg, was written in close co-operation between the experimenter and the course leaders. The design of all tests in the basic course was described in detail. The text contained 3160 words. Students supposed to have a high intrinsic motivation for reading the text were recruited from the Institute of Education itself. Students supposed to have low intrinsic motivation for reading the text were recruited from those registered for sociology, who had not earlier taken education.

Test anxiety

A couple of weeks before the experimental sessions all subjects were asked to answer a number of self-rating questions. One of the questions was intended to give a rough indication of level of test anxiety. An inventory measuring test anxiety would have given a more reliable index, but might also have aroused unwanted expectations of an anxiety-provoking experiment.

The test anxiety question is found below. The number of subjects marking each category is found after the alternatives.

How often do you get nervous in test situations?

☐ I very often get nervous in test situations. (3)

☐ I often feel nervous in test situations but most of the time I am able to control myself. (45)

☐ I get nervous only in some test situations. (22)

☐ I very seldom get nervous in test situations. (11)

The total group of subjects was dichotomised according to level of test anxiety. The number of categories on the test anxiety question turned out to be too few and too inclusive. Even the best cut put a clear majority of the subjects in the high test anxiety category, and education students reported higher levels of test anxiety than sociology students. The four experimental groups had thus to be of unequal size.

TABLE 1

FOUR EXPERIMENTAL GROUPS

	Level of intrinsic motivation	Low anxiety	High anxiety
Education	High	N=13	N=25
Sociology	Low	N=20	N=23

Experimental conditions

The subjects of each subgroup were randomly assigned to one of two experimental conditions. One of the experimental conditions was designed to eliminate as many features as possible that could create extrinsic motivation. The other condition was designed to induce rather strong extrinsic motivation. The experiment was conducted with groups of 3-5 subjects each time. In the strong extrinsic motivation condition the subjects were told that after the reading session one of the students was to be asked to report what he had gathered from the text to the rest of the group. His report was to be tape-recorded for later detailed analysis. A large tape recorder was placed in front of the subjects and prepared for immediate use. The performance session was to be finished by a group discussion on the merits of the text as information material. It was believed that by these arrangements the level of achievement in the performance session should be felt to have consequences for level of self-confidence and self-esteem. Thus it was believed that the subjects would work hard in order to defend or maybe enhance their self-esteem.

In the weak extrinsic motivation condition, measures were taken to create a friendly relaxed atmosphere. The students were instructed that after the reading session they would be asked to write down what they recalled from the text. As in the other group the session was to be finished by a general discussion on the merits of the text as information material.

In reality all subjects were treated in the same way during the performance phase. Immediately after the reading session all subjects were first asked to write summaries of the most important information of the text (test A). The question was phrased as follows: "Summarise below the most important information of the information sheet you have just read". This question was followed by an empty page.

When they had completed their summaries the subjects were asked to answer a 15 item short-answer test of factual knowledge based on the content of the text (test B). Finally each subject was interviewed individually about his way of approaching the text and his experiences while reading. The interviews were recorded and transcribed.

RESULTS

The presentation of results is divided into three sections. In the first section an analysis of ways of approaching and processing the learning material is presented. This section also includes analyses of outcomes of learning. In the second section effects of different types of motivation on process and outcome of learning are studied with subjects divided into the four experimental groups presented in Table 1. In the third section effects of different types of motivation are studied with subjects divided according to their reports of how they actually perceived the experimental conditions.

I *Qualitative differences in the process of learning*

Marton and Säljö (1976) summarised results from a series of studies conducted at the Institute of Education, Göteborg University. They found "basically two different levels of processing . . . deep-level and surface-level processing . . . In the case of surface-level processing the student directs his attention towards learning the text itself (the sign) ie, he has a reproductive conception of learning which means that he is more or less forced to keep to a rote-learning strategy. In the case of deep-level processing, on the other hand, the student is directed towards the intentional content of the learning material (what is signified) ie, he is directed towards comprehending what the author wants to say about, for instance, a certain scientific problem or principle" (pp7-8).

In spite of the fact that the learning material used in the present experiment was very different from the text used in the previous studies (Marton, 1975a; Säljö, 1975; Dahlgren, 1975; Svensson, 1976, 1977) the same two levels of processing were found when the transcribed interviews were analysed and categorised. In addition to the two levels of processing, two levels of attention were identified. Within each level of processing some subjects were working very intensively while others seemed to be satisfied by forming a general impression. The deep-level processors working with a high level of attention tried to go behind the information given in great detail. Deep-level processors working with a low level of attention tried to form a general impression of the information material in order to be able to return to the text when their need for this special information became more pressing. Surface-level processors working with a high level of attention tried to press the text into their minds, while low level of attention at this level of processing meant a kind of lazy reading hoping that at least some information might slip into the memory.

The present author first analysed all transcribed interviews a number of times, developed a classification schedule and then categorised all the students. The classification schedule was then used by another rater who again categorised all students. Complete agreement between the two classifications was reached concerning 74 per cent of the students. The final classification of the remaining subjects was made after discussions between the two raters.

The four categories are illustrated by excerpts from the transcribed interviews.

(1) Deep-level processing, high level of attention:
Subject 47
Interviewer: What were you thinking of while reading this? Were you thinking of the situation, were you thinking of the text as such or were you thinking of what the text was about, the phenomena behind it?
Student: I was primarily thinking of what it was about. I made comparisons with our own sociology course and I concluded that there were great similarities between the two courses. That was the first thing I was thinking of. Then I started to think of how it would have been if I had taken education instead of my present course.

(2) Deep-level processing. Low level of attention:
Subject 24
I: What did you think of this material you have just read?
S: Well, I found it most interesting because I don't know very much about education, you know. I didn't know anything about the instructional methods used . . . the design of the courses and their contents. It was very useful to get information about all these things.
I: What did you expect when you came here? About the material, about the situation and so on?
S: I wasn't nervous or anything like that. Maybe I became a bit nervous during the tests because I read the information sheet as I would have done at home. I know that I can read it once more if I am wondering about some detail. I just scanned the material to see what kind of information it contained.

(3) Surface-level processing, high level of attention:
Subject 43
I: Were you able to concentrate on the contents?
S: As I knew that you had to try to remember the whole thing it was very difficult to concentrate on the content—I mean in a deeper sense—what it meant. You just thought: now I must remember this and that. You tried to pick out some facts.

(4) Surface-level processing. Low level of attention:
Subject 82
I: If I have understood what you have said you were not thinking of what the text was about, the courses described, but of memorising the details?
S: Yes, I did.
I: Would you like to tell me something about how you read? Did you read it as you read a newspaper, as you read a good book or as you read course material?
S: In the beginning I read very carefully but after that I hurried through it. I lost interest, I didn't think about what I was reading.

Qualitative differences in free recall presentations

Säljö (1975) analysing answers to free recall questions identified three levels of outcome or recall strategies: mentioning, describing and conclusions-oriented answers. Students using the strategy called *mentioning* just indicate that the author discusses certain problems without commenting upon what he says or why. Säljö regards this strategy as the most superficial level or recall. The level called describing "is characterised by the subjects giving a description, often very detailed and accurate of the premises; however, they leave out the most important part, namely the conclusions that can be drawn from a certain set of facts" (p84).

Subjects using a *Conclusion-oriented* strategy "account for the intentional content or what the author wants to say in a certain context" (p84).

The text used as learning material in the present experiment differed in important respects from Säljö's texts. The authors of our text were not trying to present theoretical concepts or a set of arguments. They were just trying to give a neutral description of the system of examination of the basic course at the Institute of Education. Nevertheless the categories of answers identified by an analysis of answers to the open question of test A turned out to be very similar to the categories identified by Säljö.

The similarities are so striking that it has been natural to adopt Säljö's terminology. When the analyses presented below were carried out, however, Säljö's results were not known to this author.

After careful reading and rereading of all answers to test A descriptions of four qualitatively different categories were worked out. Two independent raters using these descriptions then categorised all students and reached agreement concerning 75 per cent of the students. The final categorisations of the remaining 25 per cent were made after discussions between the two raters. Four categories were identified:

(a) *Conclusion-oriented, content*
The student summarises his main conclusions from reading the text. He explains his thoughts and reflections while reading the text, and summarises the parts of information that he has found most interesting.

(b) *Conclusion-oriented, mentioning*
The student reports that he has found certain parts of the information sheet interesting, but he does not summarise the contents of these parts.

(c) *Description, content*
The student has tried to give a neutral and complete summary of the content of the text.

(d) *Description, mentioning*
The student has intended to write a complete list of content of the text. The answers categorised as 'mentioning' by Säljö belong to this category.

The relation between levels of processing and categories of outcome is presented in Table 2.

TABLE 2
LEVEL OF PROCESSING AND LEVEL OF ATTENTION IN RELATION TO TYPE OF FREE RECALL SUMMARY

	Deep-level processing		Surface-level processing		Sub-Totals
	High attention	Low attention	High attention	Low attention	
Conclusion-content	6	9	3	—	18
Conclusion-mentioning	3	3	3	—	9
Description-content	9	9	20	4	42
Description-mentioning	—	7	4	1	12
Sub-Totals	18	28	30	5	81

$(X^2 = 14.98, \text{d.f.} = 9, P < .10)$

There is a clear tendency for surface-level learners to try to give complete descriptive summaries. Deep-level learners distribute more evenly over the categories. A possible explanation to the pattern of results might be that deep-level processing has made it possible to choose response strategy according to the student's perception of the demands of the test situation, while surface-level processing has prepared the student for one type of performance only. Students reading at a low level of attention have a greater tendency than high-attention readers to summarise by mentioning. The relation between the processing variables and the results on the factual knowledge test (test B) is presented in Table 3.

TABLE 3

LEVEL OF PROCESSING AND LEVEL OF ATTENTION IN RELATION TO PERFORMANCE ON A FACTUAL KNOWLEDGE TEST

	Approach to learning			
	Deep-level High attention	Deep-level Low attention	Surface-level High attention	Surface-level Low attention
Mean	12.1	7.3	8.8	7.2
S.D.	3.2	2.7	3.1	1.7
N	18	28	30	5

There were reasons to expect the factual knowledge test to be adapted to the learning approach of surface-level learners reading the text at a high level of attention. The best results are not achieved by this subgroup, however, but by deep-level learners reading at a high level of attention ($t=3.46$, $P<.01$). Surface-level learners reading at a high level of attention performed only slightly better than deep-level learners reading at a low level of attention. The corresponding t-value is $t=1.93$ ($P<.10$).

II *The Experimenter's Definition of the Situation*

Students registered for education, but not subjects registered for sociology, were expected to be strongly intrinsically motivated to read the experimental text. Students given the instruction that the performance phase of the experiment would involve close inspection and assessment of their recall under socially demanding conditions were expected to be strongly extrinsically motivated for learning the text while subjects given an instruction intended to reduce ego-involvement during learning were weak in extrinsic motivation.

The analyses presented in this section take for granted that these assumptions of the experimenter were true.

In the following three tables the relations of type of motivation to level of processing, type of free-recall summary and test result are presented. The subjects are divided according to level of test anxiety.

TABLE 4
INTRINSIC AND EXTRINSIC MOTIVATION IN RELATION TO LEVEL OF PROCESSING

Extrinsic Motivation	Level of Processing	Low Test Anxiety		High Test Anxiety		Sub-Totals
		Intrinsic Motivation		Intrinsic Motivation		
		Strong	Weak	Strong	Weak	
Weak (Supportive)	Deep-level	6	9	6	7	28
	Surface-level	—	3	8	4	15
Strong (Threatening)	Deep-level	5	3	7	3	18
	Surface-level	2	5	4	9	20
Sub-totals		13	20	25	23	81

Level of extrinsic motivation is related to level of processing among students weak in intrinsic motivation. ($X^2=6.70$, $P<.01$). Extrinsically motivated students show a stronger tendency to choose surface-level processing. There is no corresponding pattern among students with strong intrinsic motivation ($X^2=.18$). This is at least partly caused by the perplexing behaviour of highly test anxious students with strong internal motivation. In this subgroup deep-level processing was more common under strong than under weak extrinsic motivation.

TABLE 5
INTRINSIC AND EXTRINSIC MOTIVATION IN RELATION TO TYPE OF FREE-RECALL SUMMARY

Extrinsic Motivation	Type of Summary	Low Test Anxiety		High Test Anxiety		Sub-Totals
		Intrinsic Motivation		Intrinsic Motivation		
		Strong	Weak	Strong	Weak	
Weak (Supportive)	Conclusion-content	1	4	4	4	13
	Conclusion-mentioning	3	—	3	—	6
	Description-content	1	6	4	7	18
	Description-mentioning	1	2	3	—	6
Strong (Threatening)	Conclusion-content	2	1	2	—	5
	Conclusion-mentioning	2	—	1	—	3
	Description-content	2	7	4	11	24
	Description-mentioning	1	—	4	1	6
Sub-totals		13	20	25	23	81

The pattern of results in Tables 4 and 5 are similar. Again it is only among students weak in intrinsic motivation that the experimental manipulation seems to have had any directing effect. ($X^2=5.73$, $P<.02$). All but one of the subjects working under weak intrinsic motivation and strong extrinsic motivation has given description-type summaries.

It is worth noting that mentioning-type summaries are given almost exclusively by students with a strong intrinsic motivation. This serves as a warning against automatically following the practice of Säljö (1975) in regarding 'mentioning' as necessarily a sign of technified (low-attention) surface-level learning.

TABLE 6

MEAN SCORES ON A TEST OF FACTUAL KNOWLEDGE IN RELATION TO DIFFERENT LEVELS OF INTRINSIC AND EXTRINSIC MOTIVATION

		Low Test Anxiety		High Test Anxiety	
		Intrinsic Motivation		Intrinsic Motivation	
Extrinsic Motivation		Strong	Weak	Strong	Weak
Weak (Supportive)	Mean	11.7	8.2	7.6	8.1
	(N)	(6)	(12)	(14)	(11)
Strong (Threatening)	Mean	11.1	9.4	9.4	8.3
	(N)	(7)	(8)	(11)	(12)

It was expected that students with strong intrinsic motivation would learn less under strong than under weak extrinsic motivation. This expectation is not confirmed. Both extrinsic motivation and intrinsic motivation are positively related to test results.

Level of test anxiety is negatively related to test achievement among students with a strong level of intrinsic motivation, but not among those with a weak level of intrinsic motivation. The measure of test anxiety indicated a general tendency to react anxiously towards test situations (test anxiety). The pattern of results is more pronounced when *state* anxiety is substituted for trait anxiety. At the end of the experimental session the students were asked to rate their level of state anxiety while they were learning the text. A questionnaire developed by Fransson (1972) was used.

Among students with strong intrinsic motivation the relation between state anxiety and test performance is very significant. $r=-0.54$, $P<.001$). Among students with weak intrinsic motivation the correlation between state anxiety and test performance is very low ($r= -0.08$). The expectation that extrinsic motivation, caused by a socially demanding learning and test situation, would negatively affect learning of intrinsically motivated subjects was based on the assumption of a close connection between extrinsic motivation and state anxiety. This assumption was obviously not correct.

III *The Subjects' Definition of the Situation*

The concluding interview revealed that the basic assumptions used in defining the experimental groups were true only to a limited extent. Some of the sociology students found the learning material of the experiment very interesting and were thus clearly intrinsically motivated. On the other hand some of the education students found the text of limited relevance. The efforts to manipulate level of extrinsic motivation also turned out to have been only moderately successful. Level of trait anxiety seemed to have been an important factor influencing how the subjects actually perceived the experimental situation. The anomalous reactions of the subgroup of education students with high level of trait anxiety became more intelligible on closer examination of their definitions of the situation. Fourteen students from this group participated in the experimental conditions designed to reduce extrinsic motivation. In spite of the attempt to create a supportive climate, nine of these students perceived the experimental situation as demanding and adapted their way of reading to an expected, but not announced, factual-knowledge test. Seven of these nine students were classified as surface-level learners. Only five out of the 14 students experienced the experimental conditions as the experimenter intended and of these four were classified as deep-level learners.

Endler (1975) reported studies showing that situational factors explained considerably less variance in emotional states among people with a high level of trait anxiety than among those with a low level of trait anxiety.

These observations formed the starting point for a second analysis based on the reports of the subjects of how they actually perceived the experimental conditions and how interesting they really found the learning material. Tables 7-9 below correspond to Tables 4-6 in the preceding section. Level of intrinsic motivation in section II corresponds to level of interest in this section. Extrinsic motivation in section II corresponds to 'adaptation to expected test demands' in this section.

TABLE 7

REPORTED INTEREST AND ADAPTATION TO TEST DEMANDS IN RELATION TO LEVEL OF PROCESSING

Adaptation to test demands	Level of Processing	Low (Trait) Test Anxiety		High (Trait) (Test) Anxiety		Sub-Totals
		Strong interest	Weak interest	Strong interest	Weak interest	
Non-adaptive	Deep-level	5	5	6	3	19
	Surface-level	—	1	—	2	3
Adaptive	Deep-level	8	5	8	6	27
	Surface-level	2	7	6	17	32
Sub-totals		15	18	20	28	81

Irrespective of their level of interest in the text, there are considerable differences in level of processing between students who have adapted their learning to expected test demands and those who have not ($\chi^2=4.75$, $P<.05$; $\chi^2=5.89$, $P<.02$). High level of trait anxiety increases the probability for surface level learning ($\chi^2=3.78$, $P<.10$). The perplexing result in Table 4 among subjects high in trait anxiety is not found in Table 7. When the analysis starts from the subjects' reported experiences, the pattern of results is no longer anomalous. Lack of interest in the text, efforts to adapt to expected test demands, and high trait anxiety, all increase the tendency towards surface-level processing.

TABLE 8

REPORTED INTEREST AND ADAPTATION TO TEST DEMANDS IN RELATION TO TYPE OF FREE-RECALL SUMMARY

Adaptation towards test demands	Type of Summary	Low (Trait) Test Anxiety		High (Trait) Test Anxiety		Sub-Totals
		Strong interest	Weak interest	Strong interest	Weak interest	
Non-adaptive	Conclusion-content	—	4	1	1	6
	Conclusion-mentioning	—	—	2	—	2
	Description content	3	1	2	3	9
	Description mentioning	2	1	1	1	5
Adaptive	Conclusion-content	3	1	5	3	12
	Conclusion-mentioning	5	—	1	1	7
	Description content	2	10	4	17	33
	Description mentioning	—	1	4	2	7
Sub-totals		15	18	20	28	81

There are no important differences between the pattern of results in Tables 5 and 8. Only among subjects with weak interest in the text is there a significant relation between adaptation to expected test demands and type of recall summary ($\chi^2=4.78$, $P<.05$). Subjects influenced by expected test demands show stronger tendency to give descriptive summaries.

TABLE 9

MEAN SCORES ON A TEST OF FACTUAL KNOWLEDGE IN RELATION TO REPORTED INTEREST AND ADAPTATION TO TEST DEMANDS

Adaptation to test demands		Low (Trait) Test Anxiety		High (Trait) Test Anxiety	
		Strong interest	Weak interest	Strong interest	Weak interest
Non-adaptive	Mean (N)	8.6 (5)	7.7 (6)	8.7 (6)	6.4 (5)
Adaptive	Mean (N)	12.4 (10)	9.0 (12)	9.2 (14)	8.0 (23)

Adaptation to expected test demands seems not to have had any destructive effects on learning among subjects strongly interested in the text. The results of Tables 6 and 9, on the contrary, indicate positive effects. In an analysis not reported here, a strong relation between efforts to adapt to test demands and intensity of learning has been found.

CONCLUSIONS

Once more it has been possible to differentiate between the qualitative different levels of processing and outcome identified previously. But it has also been possible to show that the type of motivation for reading a particular text is an important factor influencing the choice of approach to learning, and thus also determining likely levels of outcome.

A subject motivated by test demands to read a text for which he has very limited interest is very probable to adopt a surface-learning strategy, while deep-level learning seems to be the normal strategy chosen by a student motivated only by the relevance of the content of the text to his personal needs and interests.

Level of trait anxiety has been shown to be an important variable influencing the receptivity of the subject to situational cues. Expectations built up by previous experiences might sometimes be a more important factor determining the perception of the learning situation of the highly trait-anxious student than the actual situation as it is defined by the experimenter.

The conclusion to be drawn for teaching is that if deep-level processing is valued, every effort must be made to avoid threatening conditions, which rely mainly on extrinsic motivation. This is especially important when the initial interest of the students in the learning task is low.

The results support the arguments by Edfeldt (1976) and Dewey (1913) that the natural impulse of the intrinsically motivated learner, unthreatened by expectations of a factual knowledge test, is deep-level processing—a state of undisturbed contact between the author and his reader.

REFERENCES

BIGGS, J. B. (1971). *Information and Human Learning.* Glenview, Illinois: Scott Forsman.

BRUNER, J. S. (1961). The act of discovery. *Harvard educ. Rev.*, 31, 21-32.

COSTELLO, C. G. (1964). Ego-involvement, success,and failure: a review of the literature. In EYSENCK, H. J., (Ed.), *Experiments in Motivation.* London: Pergamon Press.

DAHLGREN, L. O. (1975). *Qualitative Differences in Learning as a Function of Content-Oriented Guidance.* Göteborg: Acta Universitatis Gothoburgensis.

DeCHARMS, R. (1968). *Personal Causation.* New York: Academic Press.

DEWEY, D. (1913). *Interest and Effort in Education.* Boston: Houghton Mifflin.

EDFELDT, A. E. (1976). Är det meningsfullt med studieteknisk träning? (Is there a meaningful study-skill training?). *Forskning om utbildning,* 3, 15-24.

ENDLER, N. S. (1975). A person-situation interaction model for anxiety. In SPIELBERGER, C. D., and SARASON, I. A. (Eds.), *Stress and Anxiety.* New York: Halstead Press.

FRANSSON, A. (1972). *Testängslighet och språkfardighet* (Test anxiety and language proficiency). Göteborgs universitet: Pedagogiska Institutionen.

LEFCOURT, H. M. (1976). *Locus of Control: Current Trends in Theory and Research.* New York: Wiley.

LEKARCZYK, K. T., and HILL, K. T. (1969). Self-esteem, test anxiety, stress and verbal learning. *Dev. Psychol.,* 1, 147-154.

MARTON, F. (1974). Some effects of content-neutral instructions on non-verbatim learning in a neutral setting. *Scand. J. educ. Res.,* 18, 199-208.

MARTON, F. (1975a). *On Non-Verbatim Learning: I—Level of Processing and Level of Outcome.* Report from the Institute of Education, University of Göteborg, No. 39.

MARTON, F. (1975b). *On Non-Verbatim Learning: II—The Erosion Effect of a Task-Induced Learning Algorithm.* Reports from the Institute of Education, University of Göteborg, No. 40.

MARTON, F., and SÄLJÖ, R. (1976a). On qualitative differences in learning. I—Outcome and process. *Br. J. educ. Psychol.,* 46, 4-11.

MARTON, F., and SÄLJÖ, R. (1976b). On qualitative differences in learning. II—Outcome as a function of the learner's conception of the task. *Br. J. educ. educ. Psychol.*, 46, 115-127.

PARKNÄS, L. (1976). Ännu mer konkret om motivation I. (Still more concrete about motivation I). *Nordisk psykologi*, 28.

SALTZ, E. (1971). *The Cognitive Bases of Human Learning.* Homewood, Illinois: Dorsey Press.

SARASON, I. A. (1975). Test anxiety, attention and the general problem of anxiety. In SPIELBERGER, C. D., and SARASON, I. A., *Stress and Anxiety.* New York: Halstead Press.

SVENSSON, L. (1976). *Study Skill and Learning.* Göteborg: Acta Universitatis Gothoburgensis.

SVENSSON, L. (1977). On qualitative differences in learning: III—Study skill and learning. *Br. J. educ. Psychol.*, 47, November 1977.

SÄLJÖ, R. (1975). *Qualitative Differences in Learning as a Function of the Learner's Conception of the Task.* Göteborg: Acta Universitatis Gothoburgensis.

WHITE, R. W. (1959). Motivation reconsidered: the concept of competence. *Psych. Rev.*, 66, 297-333.

(Manuscript received March 7, 1977)

Br. J. educ. Psychol., 47, 258-267, 1977

ACHIEVEMENT AND SELF-REPORTS OF RESPONSIBILITY FOR ACHIEVEMENT IN INFORMAL (OPEN-SPACE) AND TRADITIONAL CLASSROOMS

By ANNE E. BELL, D. S. ABRAHAMSON
and ROSEMARY GROWSE

(Children's Centre, Winnipeg, Canada)

SUMMARY. Multivariate analysis showed that subsequent to the third year in a longitudinal study, 40 children chosen at random from traditional classrooms were superior in reading and math problem-solving, but less willing to accept responsibility for negative intellectual events, when compared with 40 children from an informal, open-space school. The groups had shown no significant differences due to type of school or sex of subject on a pre-school predictive test battery. Differences were attributable only to type of school at the fourth year. When stepwise discriminant analysis was employed with these groups only problem-solving and concepts in mathematics entered the multiple regression equation as significant predictors of group membership.

INTRODUCTION

Incidental learning and the freedom of the child to set his own learning pace are primary concepts in the philosophy of ungraded or non-graded, informal, and/or progressive educational programmes. The term open-space is often used synonymously with these since the open-space school structure would appear to facilitate the movement of students from one level of achievement to another, to give the student a greater freedom within the classroom space, as, for example, in being able to move at will from one interest to another or to and from the library. Less supervision by an authoritarian figure, a more permissive attitude toward the use of free time, self-initiated learning on the part of the child, and a recognition of individual differences and needs by the teacher are, ideally, characteristics of the open-space, non-graded programmes (Goodlad and Anderson, 1963; Tewksbury, 1967).

The traditional graded classroom on the other hand is often seen as thwarting such needs having become associated with a 'pass-fail' philosophy. According to Sobel and Tejirian (1972-3) the child must choose between expressing fundamental needs and being 'bad' in the teacher's and his own eyes, or suppress his needs in order to be 'good', thus creating emotional conflicts. Children are said to have basic needs to experience success, to feel competent and exercise autonomy. Traditional classrooms, therefore, are thought to create anxiety through their grading orientation and punishment of errors. The teacher in a traditional classroom may be regarded by the child as a punitive person because she must place restrictions on his behaviour and often appear to be a negative evaluator. The teacher in a more permissive situation where the child is to be trusted to develop, is thought more likely to create a learning climate conducive to social and emotional health for the learner (Rogers, 1969). Ideally the child's attitudes toward school and future learning should be positively oriented and his relationships with teachers and peers should reflect openness and warmth (Sobel and Tejirian, 1972-3). It is assumed that academic achievement will be optimal under such favourable conditions.

Ward and Barcher (1975) and Katz (1973) have noted, however, that the assessment of the products of open education is not being rigorously pursued, that evaluation is largely dependent on personal testimony rather than upon scientific investigation. Furthermore, they suggest that studies should recognise that variables such as age and sex may interact in a complex fashion with IQ level and the degree of 'openness', in the assessment of products, whether of achievement or an aspect of personality. Gardner (1966) in a British study of informal education reports reading results which were dependent upon age differences. Young children in informal classrooms had lower scores than those in traditional classes while older children in informal situations were superior. In contrast a more recent British study (Bennett, 1976) showed that 11-year-old children in formal classrooms had made from three to five months more progress than their age peers who had been in informal classrooms. Ward and Barcher (1975) have suggested that, as regards reading, children taught in informal classrooms may eventually arrive at the same level of achievement though moving at a slower pace. Longitudinal studies, they say, could provide some confirmation of this hypothesis by identifying such developmental patterns.

Inasmuch as the informal school programme is considered by its proponents to provide an optimal environment for the social and motivational development of the child, questions have been raised regarding the emergence of such traits as creativity, leadership, initiative and independence outcomes. It seems reasonable to assume that a child's situation would be less anxiety-provoking in the more permissive open-space school than in a traditional class. He should feel more free to fail if there has been less arousal of guilt feelings associated with lack of achievement relative to his peer group.

The major intention of the present study is to report whether or not there are differences in both academic achievement and the assumption of responsibility for academic success or failure as a function of open *versus* traditional classroom experience at the end of the third academic year. Coincidental with the major intention, the contribution of sex of pupil to differences in the above variables is to be assessed. Again conditional to the demonstration of differences, the identification of measures which singly and in combination contribute to such discrimination is to be made.

The present project has some unique aspects in comparison with previous studies regarding personality characteristics, for example, Ramey and Piper (1974), Reiss and Dyhdalo(1975), and Ward and Barcher (1975), in that being longitudinal in design, pre-school information regarding intellectual and social status about the groups is available. Additionally, each child is known to have spent four years in either a closed or open programme, and his progress through the years has been followed closely (Bell *et al,* 1976). Cross-sectional methods, even with optimal matching procedures, for example, age, sex and IQ at the fourth grade, leave open to doubt such variables as the school readiness of the child when he entered school and the actual length of time he has been in an open-space or classroom situation.

For the present study the children in the two groups at school entry did not differ significantly in the skills measured by a pre-school battery, nor did they differ in socio-economic status. Also Bell *et al* (1973, 1974) have shown that the first two years were critical for the total population of beginners in a single year in these two schools with regard to the acquisition of basic academic skills. This finding which is similar to that of Gardner (1966) in the early years in British schools, makes still more questionable the choice of a child at the third or fourth year as representative of either traditional or open-space children, respectively, solely because he is presently found in that setting.

METHOD

Schools

Both schools involved in this study were located in a middle-class suburban area, roughly a mile apart. The children, therefore, shared a community environment with regard to cultural, experiential and socio-economic variables. Placement of the child in either of the schools was arbitrary since school areas were defined and there was no choice of programme within a school. The school designated as traditional (TC) differed in structure from the open-space (OS) school in that each of four first-grade classes had its own enclosed classroom with its own teacher. The beginners in the OS school (N=57) shared a large open space with the total enrolment which included ages 5.7 to 12 years (N=350 students). Sliding walls or dividers could be moved at will but were seldom used.

Teachers in the OS programme had chosen their role and had received additional training in the philosophy and operation of OS programmes. They, together with those in the traditional classrooms, were well-trained and experienced in the handling of primary children. The two primary teachers in the OS school had 28 and 29 students, respectively, in their groups. These groups were similar in size to those within each of the four classrooms in the larger TC school where 112 beginners had been admitted. The classes in the OS school were referred to as years rather than grades, and subdivisions within years were termed levels. The four TC beginner classrooms were known as grades but the expected first-year work in reading in both schools was the mastery of three levels of achievement using the same basal tests, viz, Level 1 (primers), Level 2 (first reader), Level 3 (second reader). Second- and third-year (grade) work included two further levels within each year as the norm.

Sample

At the beginning of the fourth academic year 86/112 children remained in the TC and 48/57 in the OS school, 20 of the latter being girls. Matching in numbers for sex thus placed a limitation on the size of the available groups. An investigator who had no previous knowledge about the children, selected at random 20 boys from the 28 left in the OS school. Two similar groups, 20 boys and 20 girls were selected at random from the 86 children in the TC school. All of these children would be nine years old before the end of the fourth year. There had been no significant differences in mean chronological age, either between schools, or between boys and girls within a school when they were beginners. Attrition over the years had not been selective, therefore no significant differences in these variables occurred at the fourth year. The Blishen Index (1967), which is a numerical rating given on the basis of the father's occupational status, showed that the four groups did not differ significantly in economic background.

Procedure

Pre-school Battery. As a preliminary to the longitudinal study all of the children in their kindergarten year had been tested on variables generally held to be related to learning skills. These included general intelligence, visual-motor co-ordination, visual perception, auditory memory and discrimination, and a school readiness screening test. The standard tests used were the Stanford-Binet Intelligence Scale (Terman and Merrill, 1960), Bender-Gestalt (1938), Visual Organisation Test (VOT; Hooper, 1957), and the School Readiness section of the Early Detection Inventory (EDSR; McGahan and McGahan, 1967). The Bender-Gestalt (1938) was scored by the Koppitz (1964) method, but with errors deducted from the total possible (30) to obtain a success score which would be compatible with the ranking system of the other variables. The auditory test which included memory for both sentences and digits, also measured the child's ability to discriminate consonants and vowels, as well as blend sounds. It was administered individually by means of a tape.

The School Readiness section of the EDSR was scored numerically by giving 2, 1, or 0 points to the categories designated as successful, partially correct, and unsuccessful, by test norms. Questions are based on four aspects of child behaviour which are examined by means of 11 subtests. They concern integration of body parts or self-image, orientation in the environment, and ability to perceptualise as shown by awareness of left and right. In this measure, concept level is assessed by an ability to recognise basic colours and coins, by counting up to 13 coins and showing an understanding of number beyond rote counting. Fine motor co-ordination is measured by the copying of simple figures and the drawing of a person.

Achievement Tests. The tests used over a period of three years to measure progress in reading and number were: Schonell Word Recognition (1961), an individually administered word list not related to vocabulary of any specific reading series; Gates-MacGinitie Vocabulary A, B, and C, Form 1 (1965), and Co-operative Primary Number Concept, 12A and 23A (1967). The latter group test is concerned with such concepts as more or less, odd or even, sequencing of numbers (16, 12, ?, 4), estimations of amount, length, etc. No computation is required. The Canadian Test of Basic Skills (King, 1967) was administered at the end of the third year. All of the latter were group paper and pencil tests.

The Intellectual Achievement Responsibility Scale. This measure, given in the early spring of the fourth school year, attempts to measure beliefs in internal *versus* external control. External forces in this scale are the persons who most often come into close contact with the child, and from whom he must seek to resolve his dependency as a natural progression in development, ie, his parents, teachers and peers. No assertions as to the operation of luck or fate are involved. This scale samples an equal number of positive (IAR+) and negative (IAR−) events since the authors (Crandall *et al,* 1965) believe that responsibility for success and failure might develop at different rates in young children, or, for some children and not for others. Test-retest reliabilities for grades three to five are reported by the authors to be .66, (IAR+) and .74 (IAR−). Internal consistency for the same grades is reported to be .54 for IAR+ and .57 for IAR−. In addition, correlations between IAR+ and IAR− are said to be so low for young children (eg, Grade 4, $r=.11$) as to raise doubts about the value of the total IAR score as an experimental variable.

Administration. Attrition in the original numbers occurred yearly, while newcomers joined the research children in their classes. The achievement tests were administered to all the children in a classroom, or group, so that the research children were not singled out. As the children moved through the grades, new teachers would not necessarily know which children were being followed. The reported results, however, concern only those who had attended the same school from kindergarten to the fourth year and participated in the pre-school testing procedures.

Statistical Analyses

In an attempt to provide some control for error rate in an experimental design (Kirk, 1968), assessment of differences between type of school and sex of child in pre-school information was made by means of a 2×2 (OS *vs* TC) (M *vs* F) multivariate analysis of variance (Manova). A similar 2×2 Manova was applied to the three achievement variables and the two responsibility measures. To clarify the nature of possible differences in these variables due to type of school and/or sex of the subject, a stepwise discriminant analysis was planned.

RESULTS

Pre-school Battery

The two original groups of beginners (OS=57, TC=112) had shown no significant differences in mean scores on any of the predictor variables or on the socio-economic index (45.9 and 44.5 respectively) when they were tested at the end of the kindergarten year. The groups, therefore, may be considered as well-matched on IQ (OS=113.9, TC=112.1) and on auditory, visual, and visual-motor skills. The screening test for school readiness (EDSR) gave them equivalent ratings (OS=19.3, TC=18.8).

Achievement: Years One and Two

Both individual (Schonell) and group reading (Gates-MacGinitie) mean scores were significantly better for TC children at the end of the first year. They had a seven months' advantage over the OS students (Sch. $t=5.012$, $P<.001$; G-M. $t=4.007$, $P<.001$).

Reading results at the end of the second year showed that the TC classes were still significantly better readers (Sch. $t=2.963$, $P<.01$; G-M. $t=2.899$, $P<.01$) by seven and six months, respectively. Number Concepts showed no significant difference at either year.

Achievement and IAR: Years Three and Four

The Manova (BMD12V; Dixon, 1973) was first applied to the seven pre-school variables. No significant differences due to type of school, sex of child, or their interaction were found. Mean IQ scores for the two groups were 113.29 and 114.73 for OS and TC respectively, a bright normal level. The Blishen Indices of 43.56 (OS) and 41.54 (TC) showed that the groups were well-matched in terms of socio-economic status (middle-class). Thus the mean scores of the groups (N=40) randomly selected from the students remaining at the end of the third year closely matched those of the entire groups when they began the school programme.

A similar Manova performed on the third-year achievement tests (Gates-MacGinitie, CTBS Math Concept, CTBS Problem Solving), and fourth-year responsibility variables yielded significance only for type of school $F(5,72)=6.243$, $P<.001$.

TABLE 1

DATA ON THE ACHIEVEMENT AND IAR MEASURES
FOR OS AND TC SCHOOLS AND THE NORMATIVE SAMPLES

	Norms		OS		TC	
	Mean	SD	Mean	SD	Mean	SD
G-M Reading	35.00	9.82	35.50	9.35	39.45	7.45
CTBS Maths Concepts	3.9	1.04	4.2	1.14	4.3	0.96
CTBS Prob. Solving	3.9	1.09	4.0	1.12	4.7	0.88
IAR+	12.51	2.13	13.38	1.79	12.78	1.91
IAR−	12.26	2.35	12.03	2.26	10.90	2.61

The Gates-MacGinitie raw scores have grade equivalents of 3.9 for OS and 4.6 for TC. Interpretation of grade scores is: 3.9, late third year; 4.6, six-tenths of the way through fourth year requirements. CTBS scores are already given as grade scores.

Means, standard deviations, and test norms for the achievement and responsibility scores are shown in Table 1. Significant differences between schools occur on the reading test (G-M) $F(1,78)=4.370$, $P<.05$; on CTBS Problem Solving $F(1,78)=8.890$, $P<.005$; and IAR- $F(1,78)=4.250$, $P<.05$. While beginning the first year with pre-school skills closely comparable as shown by the first Manova, the TC children still have a seven months' advantage in reading as did the entire group through the first and second years. The difference in number concept though measured by a different test (CTBS) remains non-significant. Significant differences in mathematical problem solving from the CTBS favoured the TC children also by seven months, while the OS group showed significantly better scores for IAR-. Since no significant sex differences in performance were found, a stepwise discriminant function analysis was performed in order to identify a subset of the variables which 'best' discriminated between the two populations, OS (N=40), TC (N=40). In line with the procedure recommended by Kirk (1968, pp252-3) type of classroom was the criterion and the five achievement and responsibility scores were entered as the predictor variables (BMDP2R; Dixon, 1975).

TABLE 2

INTERCORRELATIONS OF ACHIEVEMENT AND RESPONSIBILITY MEASURES

	Sch.	Reading G-M	Maths Concepts	Maths Problems	Resps. IAR+
School*					
Gates-MacGinitie	—.230				
CTBS Maths Concepts	—.026	.746			
CTBS Prob Solving	—.320	.771	.779		
IAR+	.162	.064	—.088	—.045	
IAR—	.227	—.157	—.041	—.128	.012

* Traditional classroom coded as —1.

The zero-order intercorrelations of achievement and responsibility measures with type of school are shown in Table 2. The discriminant analysis, ($\propto \simeq .05$), indicated that the variable for which mean values in the two samples were 'most different' was CTBS Problem Solving ($F(1,78)=8.889$). The second entry was Number Concept ($F(1,77)=12.638$). The stepwise process stopped at this point, indicating that none of the other variables significantly contributed to the discrimination between the two groups. Number Concept shows no zero-order correlation with school membership, but does show a significant zero-order correlation with CTBS Problem Solving (.779, $P<.001$), suggesting that it has acted as a suppressor variable. That is, after the entry of the two CTBS variables into the multiple regression equation, neither G-M reading nor IAR- carry enough unique predictive variance to add significantly to the prediction of the school in which a child should be located on the basis of the achievement and responsibility scores.

DISCUSSION

While the open-space children showed a significantly higher mean score on IAR-, ie, the acceptance of responsibility for their failures, than did the graded children, the actual functional significance of this finding remains in some doubt in view of the fact that the mean scores for both groups are well within normal range bounds (±1 *SD)* for the standardisation sample (see Table 1, Crandall *et al,* 1965). In addition, there is the evidence from the discriminant analysis that the CTBS variables (Problem Solving and Maths Concepts) are sufficiently predictive of the OS-TC differences and are related to Gates-MacGinitie (reading) and to IAR- scores (responsibility for failure) to a degree adequate enough to carry the load of predictive variance. Nevertheless, the tendency to be less accepting of responsibility for failures shown by the traditional class children may be indicative of a somewhat higher anxiety level in comparison with that shown by children in the more permissive open-space situation. Previously published data on these groups (Bell *et al,* 1976) would seem to support such a view. As in the case of the IAR- scores, the anxiety level was, however, within normal bounds. The authors questioned the extent to which the heightened anxiety level may have contributed to the better achievement found in the graded school.

The appearance of a suppressor variable was unexpected but could have been perhaps anticipated from an inspection of the original correlation matrix between all variables of the study (see Table 2). A suppressor variable, by definition, is a predictor (in this case, CTBS Maths Concepts) which fails to correlate at a zero-order level with the criterion variable (type of school), but correlates substantially with another predictor variable (CTBS Problem Solving). Its contribution to a multiple regression equation traditionally is interpreted as the elimination of error variance from the other predictor-criterion relationships (Van de Geer, 1971). An explanation of this variable obviously must involve some characterisation of the variance that is shared by CTBS Problem Solving and CTBS Maths Concepts, and a rationale for the failure of this variance to show relevance to the criterion. The correlation obtained between the two CTBS variables indicates that approximately 60 per cent of their variance is shared. CTBS Maths Concepts, as previously described, assesses what may be called the common language of numbers. That is, beyond the actual reading of the question, the child is required to respond with knowledge that may be gained incidentally through experience and critical observation and is probably more highly related to intelligence levels and awareness than to the factual teaching of mathematics which is required for accurate problem solutions. In addition, the Maths Concept test is somewhat less demanding of reading skills in that pictorial cues are available. Open-space and traditional groups were not differentiated either in intelligence level or their scores on Maths Concept tests from the first grade on, as the present scores and previous data have shown (Bell and Switzer, 1973). While such knowledge and adequate levels of IQ are basic to success in CTBS Problem Solving, the tasks of this subtest are more demanding in terms of additional skills involved. First, the child must be able to read the problem accurately and with comprehension. Second, he must determine which arithmetic process (eg, addition, subtraction, division) will be required to solve the problem, and third, he must have learned basic number facts. This last skill in particular presupposes drill, which being repetitious, boring, and suggestive of failure where the slower learner is concerned is, theoretically, not a feature of the open-space programme. Thus while the Maths Concept subtest fails, of itself, to differentiate between the two groups because of shared variance related mainly to intelligence levels, the significantly better work of the graded children on the Maths Problem Solving subtest may be due to some unique variance associated with the more practical and structured aspects of the school programme.

A question which may be raised is the extent to which achievement scores reflect the influence of the type of structure on the carrying out of the programme, or to some misinterpretation of the needs and capacities of the beginners. Pertinent to the latter, Packard (1972-3) for example, has suggested that young children often seem to seek direction, lacking the skills which would permit them to make productive use of such freedom as that afforded by the open-space environment. Williams (1971) has questioned the wisdom of bringing untrained, asocial, and/or uncommunicative children into the more permissive situation of the open-space school. In other words, are most beginners behaviourally ready for self-regulated learning? Relating to both structure and the needs of the child, Reiss and Dyhdalo (1975) have noted that the noise level and other distractions of the open situation mitigated against optimal learning conditions, particularly in the case of non-persistent or distractible boys. Along these lines a paper concerning the present children at the first year (Bell *et al,* 1974, p415) states that "the typical child who makes a reasonable success in the open-area may be described as showing at least normal range perceptual, motor, and cognitive skills. In addition, his social behaviour is appropriate to his chronological age, and he gives some indication that he has already acquired a learning set, having mastered a number of basic concepts. He must be able to centre his attention in spite of a confusion of distracting stimuli." The proportion of such children would be roughly the same in both school populations but the failure rate in the OS school more than doubled that in the TC school (28 per cent *vs* 13 per cent). It seems, therefore, that the more sheltered, less distracting situation of the closed classroom had been more effective than the OS environment in bringing along the children less ideally equipped in basic skills and temperament at school entry. As Reiss and Dyhdalo (1975, p512) have said, "the ideal format might provide both open-space and conventional options to meet the needs of different types of children". This would provide time for the gathering of factual materials about the possible long-range benefits of open-space programmes while excluding "the possibility of substantial harm to certain types of children".

The children in the present study became well-known as individuals to the workers who visited the school at least weekly over a four-year period. Though observations were not systematically recorded the impression gained was that the children not self-motivated and/or lacking essential skills and interest wasted much of their time in aimless wandering about, in watching the movements of other classes and interacting without useful purpose with their own classmates. How much incidental or self-initiated learning was actually taking place would be difficult if not impossible to estimate. Certainly poor study habits and attitudes towards concentrated work were in the process of formation. Further, in comparison with the traditional classroom teachers, those in the open school were required to expend considerable unproductive energy when trying to engage the attention of the whole class for any purpose, thus the more mature, teachable children were often deprived of maximum stimulation.

The authors make no claim that the results reported in this paper may be generalised widely for several reasons. First, only two schools are involved. The achievement results were, however, replicated by a second group entering these schools the following year with different teachers. Second, while the importance of the teacher variable cannot be denied, there was no reason to believe that teacher training, teaching method and/or philosophy regarding individual differences was a confounding variable in this case. Over a period of five years including kindergarten, these children came in contact with many teachers, and class formations differed each year. Third, these results are limited to the primary years and, therefore, cannot be interpreted as a denunciation of open-space innovations for older children. Within its limitations, however, it does suggest that the open-space school is less successful in equipping many young children with the basic tools of further learning. The deficit in reading and the computational aspects of mathematics still existed at least to the end of the third academic year, a partial answer to the question raised by Ward and Barcher (1975) with regard to an age at which open-space children might be found to 'catch up' with traditional class peers in achievement. The results with regard to the acceptance of responsibility are equivocal. Further investigation through follow-up of these children may provide clarification.

ACKNOWLEDGEMENTS—The authors wish to acknowledge with thanks the contributions made by Dr K. N. McRae, Director, and Fred Switzer, psychologist of the Child Development Clinic, Health Sciences Centre, Winnipeg, Manitoba, to the longitudinal research programme of which this study forms a part. They are further indebted to Dr L. J. Breen, University of Manitoba, for assistance in the preparation of this manuscript. Thanks are due also to Dr P. Coleman, Superintendent of St Boniface Schools, and, not least, the principals and teachers, the children and their parents, without whose co-operation the study could not have progressed.

Continuing financial support was provided by the Children's Hospital Research Foundation Inc., and the Manitoba Medical Services Foundation, Inc. A Public Health Research Grant (Project No. 606-7-193) of the National Health Grants programme, Ottawa, Canada, partially supported the initial stages of the study.

REFERENCES

BELL, A. E., and SWITZER, F. (1973). Factors related to pre-school prediction of academic achievement: beginning reading in open-areas *vs.* traditional classroom systems. *Manitoba J. Educ.,* 8, 21-27.

BELL, A. E., SWITZER, F., and ZIPURSKY, M. A. (1974). Open-area education: An advantage or disadvantage for beginners? *Percept. mot. Skills,* 39, 407-416.

BELL, A. E., ZIPURSKY, M. A., and SWITZER, F. (1976). Informal education in relation to achievement and personality. *Br. J. educ. Psychol.,* 46, 235-243.

BENDER, L. (1938). A visual-motor Gestalt test and its clinical use. *Am. Orthopsychiat. Ass.,* Monogr., 3.

BENNETT, S. N. (1976). *Teaching Styles and Pupil Progress.* London: Open Books.

BLISHEN, B. R. (1967). A socio-economic index for occupations in Canada. *Can. Rev. Socio. Anthropol.,* 4, 41-53.

Co-operative Primary Tests. (1967). Princeton, New Jersey: Educational Testing Service.

CRANDALL, V. V., KATKOVSKY, W., and CRANDALL, V. J. (1965). Children's beliefs in their own control of reinforcements in intellectual academic achievement situations. *Child Dev.,* 36, 91-109.

DIXON, W. J. (1973). *BMD: Biomedical Computer Programs.* Los Angeles, University of California Press.

DIXON, W. J. (1975). *BMDP: Biomedical Computer Programs.* Los Angeles, University of California Press.

GARDNER, D. E. M. (1966). *Experiment and Tradition in Primary Schools.* London: Methuen.

Gates-MacGinitie Reading Tests. (1965). New York: Teachers' College Press, Teachers College, Columbia University.

GOODLAD, J. I., and ANDERSON, R. H. (1963). *The Non-graded Elementary School.* New York: Harcourt, Brace and World.

HOOPER, H. E. (1957). *The Visual-Organization Test.* California: Western Psychological Services.

KATZ, L. G. (1973). Research on open education: problems and issues. In HEARN, D. D., BURDIN, J., and KATZ, L. (Eds.), *Current Research and Perspectives in Open Education.* Washington, D.C. American Association of Elementary, Kindergarten, and Nursery Educators.

KING, E. M. (Ed.), (1967). *Canadian Tests of Basic Skills.* Don Mills, Ontario: Nelson.

KIRK, R. E. (1968). *Experimental Design: Procedures for the Behavioral Sciences.* Belmont, California: Brooks-Cole.

KOPPITZ, E. M. (1964). *The Bender-Gestalt Test for Young Children.* New York: Grune and Stratton.

MCGAHAN, F. E., and MCGAHAN, C. (1967). *The Early Detection Inventory.* Chicago: Follett Educational Corporation.

PACKARD, R. G. (1972-3). Do we have to do what we want to do today? Structure in an open classroom. *Teachers Coll. Rec.,* 74, 553-557.

RAMEY, C. T., and PIPER, V. (1974). Creativity in open and traditional classrooms. *Child Dev.,* 45, 557-560.

REISS, S., and DYHDALO, N. (1975). Persistence, achievement, and open-space environments. *J. educ. Psychol.,* 67, 506-512.

ROGERS, C. (1969). *Freedom to Learn.* Columbus, Ohio: Merrill.

SCHONELL, F. (1961). *The Psychology and Teaching of Reading.* London: Oliver and Boyd.

SOBEL, H. W., and TEJIRIAN, E. (1972-3). The case for open education. *Teachers Coll. Rec.,* 74, 559-565.

TERMAN, L. M., and MERRILL, M. A. (1960). *The Stanford-Binet Intelligence Scale.* (3rd Rev.), Boston: Houghton-Mifflin.

TEWKSBURY, J. L. (1967). *Non-grading in the Elementary School.* Columbus, Ohio: Merrill.

VAN DE GEER, J. P. (1971). *Introduction to Multivariate Analysis for the Social Sciences.* San Francisco: W. H. Freeman.

WARD, W. W., and BARCHER, P. R. (1975). Reading achievement and creativity as related to open classroom experience. *J. educ. Psychol.,* 67, 683-691.

WILLIAMS, T. B. (1971). Too much freedom? *The Reading Teacher,* 25, 2, 177.

(Manuscript received April 27, 1977)

Br. J. educ. Psychol., 47, 268-273, 1977

THE RELATION OF SPELLING ERRORS TO COGNITIVE VARIABLES AND WORD TYPE

By JUDITH D. GOYEN and M. MARTIN

(Macquarie University, Sydney, Australia)

SUMMARY. With one exception, all of the phonetic and non-phonetic spelling errors of a sample of 93 13- and 14-year-old boys loaded significantly on a general factor, described as verbal-intellectual ability. Auditory sequential memory had a moderate loading on this factor, but visual sequential memory did not load significantly on any of the six factors. Analyses of the relation of word type to spelling performance indicated that spelling difficulty was a function of word frequency but not of regularity of orthography.

INTRODUCTION

Simon and Simon (1973) conjectured that the spelling errors of a minority of students were mainly non-phonetic in nature, and that these students probably required different instructional techniques from the majority of students whose spelling errors were mainly phonetic. The main purpose of this study was to determine whether these two types of error were independent of each other, and if so, whether a preponderance of either type was associated with different cognitive abilities. It was recognised that different methods of spelling instruction may result in different kinds of spelling errors (Peters, 1967), but an investigation of this relationship was beyond the scope of the present study.

The cognitive variables that have been studied most frequently in relation to spelling achievement are visual and auditory memory (Bannatyne and Wichiarajote, 1969; Day andWedell, 1972; Naidoo, 1972; Reid and Hughes, 1974), intelligence (Holmes, 1954; Lyle, 1969; Nelson and Warrington, 1974; Reid and Hughes, 1974), and reading (Lyle and Goyen, 1969; Reid and Hughes, 1974; Rudisill, 1957; Townsend, 1947). The results of these studies are not all in agreement, and in only two of the studies (Day and Wedell, 1972; Nelson and Warrington, 1974) were cognitive variables related to specific types of spelling errors.

Day and Wedell (1972) compared three groups of children: one with relative strengths in visual sequential memory, one with relative strengths in auditory sequential memory, and one with equal strengths in visual and auditory memory. There were no significant differences between the three groups in overall spelling achievement, but certain types of spelling errors did differentiate between groups. Phonetic accuracy was not one of the discriminating types of error. In Nelson and Warrington's (1974) study, however, phonetic accuracy did discriminate between groups. Pupils retarded only in spelling made significantly more phonetically accurate errors than those retarded in both spelling and reading, while the latter made correspondingly more phonetically inaccurate (non-phonetic) errors.

Studies of the relationship between word-type and spelling achievement suggest that spelling difficulty is a function of either word frequency or grapheme-phoneme regularity (Bloomer, 1961, 1964; Brown, 1970; Cramer, 1970; Lester, 1964). Of these investigators, only Brown (1970) and Lester (1964) examined the interaction between frequency and regularity. Lester (1964) claimed that the spelling errors of American college students clustered in the high frequency-low grapheme-phoneme regularity category, but cited no empirical evidence. Brown (1970), on the other hand, found that college students were relatively poor spellers of low frequency-high regularity words. The relation of frequency and regularity to the spelling performance of younger students is not known.

The aim of the present study was to relate the spelling errors of secondary school students to visual and auditory sequential memory, intelligence, reading, and writing speed. The relation of spelling performance to the frequency and regularity of words was also to be examined.

METHOD

Sample

Ninety-three boys aged 13 to 14 years were selected from the second form (eighth year) of a state secondary school in a middle-class suburb of Sydney. All second form boys were tested except those who could not speak English and those who were absent on all testing occasions. Girls were not included in the sample as they made significantly fewer spelling errors than the boys, and it was errors that were being investigated.

Tests

(1) *Spelling*

The spelling list comprised 50 words which were classified according to the frequency of their occurrence and the regularity of their orthography: (a) frequent-regular (10 words), (b) infrequent-regular (15 words), (c) frequent-irregular (10 words), (d) infrequent-irregular (15 words).

Frequency was derived from the word counts listed in Part I of Thorndike and Lorge's (1968) word book. A word was defined as frequent if it occurred more than 1,000 times per million words in any of the five word counts. An infrequent word was one which occurred fewer than 50 times per million words in all five word counts.

Regularity was based on Lee's (1960) method which ranked the frequency with which graphemes represented each phoneme. Although this method did not consider either the morphemic elements of words or the position of letters within words, its advantage was that it yielded a continuous score from 1.0 (highly regular) to 6.0 or more (highly irregular). In Lee's (1960) study of reading, a score of 1.6 separated the regular words from the irregular words. In the present study, words with scores of 1.7 to 2.4 inclusive were excluded from the spelling list in order to avoid the problem of classifying words with borderline scores. A word was classified as regular, therefore, if it obtained a score of 1.6 or less, and irregular, if it obtained a score of 2.5 or more.

The 50 words in the spelling list were selected from an original sample of 120 words which were pretested on a group of 54 boys who were not involved in the study. All homonyms and non-discriminating words were excluded from the spelling list. The final test had a correlation of .93 with the Schonell graded spelling test, and a test-retest reliability of .97.

For the spelling test, the order of the 50 words was randomised and each word was embedded in a sentence so that the presentation took the usual form of word, sentence and word. The spelling test was recorded by an experienced radio announcer to control for presentation rate and pronunciation. Subjects were required to write the words on special answer sheets. The total number of words spelt incorrectly by each subject was tabulated for each of the four categories of words.

Classification of Spelling Errors. The system used to classify spelling errors was adapted from that of Spache (1940). There were two main classes of errors: phonetic (phonetically accurate) and non-phonetic (phonetically inaccurate), and each class was subdivided into four sub-groups: omissions, additions, transpositions and substitutions. Every incorrect grapheme was classified according to this system. In addition, there were two classes of errors for words which were incomplete or unrecognisable. For each subject, therefore, there were 10 classes of spelling errors. The inter-scorer reliability of the entire classification based on number of errors was .99, which compares favourably with Spache's (1940) reliability of .95.

(2) *Memory*

(a) *Visual Memory.* A visual sequential memory test was developed for group presentation. This test comprised 10 different shapes which were difficult to label but easy to reproduce. The shapes were arranged in sequences of one to five, and there were three items at each level of difficulty. Each item was projected on to a screen for 10 seconds and subjects were required to reproduce the shapes from memory. The test-retest reliability was .76.

(b) *Auditory Memory.* The auditory sequential memory test of the ITPA was adapted for group presentation. Digits were presented by audio-tape at a rate of two digits per second, and subjects were required to write down the digits from memory. The test-retest reliability was .84.

(3) *Intelligence*

The Standard Progressive Matrices Test (Raven, 1938) was administered as it does not involve reading, and is generally considered to be a good measure of non-verbal intelligence. The test-retest reliability, according to the test manual, is .75.

(4) *Reading*

In order to obtain a range of reading behaviours, two standardised reading tests were administered: (a) the ACER Silent Reading Test, Form D, which provided separate scores for speed and accuracy; and (b) the Progressive Achievement Test, Form A, which provided separate scores for comprehension and vocabulary. According to the test manuals, the reliability coefficients of both tests are above .80.

(5) *Speed*

Since spelling achievement may also be related to speed of writing, two speed tests were included: (a) the Coding subtest of the WISC, which is known to be related to reading achievement (Lyle and Goyen, 1969); and (b) Writing Speed, a test in which subjects were required to write as many crosses as possible in a grid of squares within a two-minute time limit.

Procedure

All subjects were tested during school hours by the same examiner and one invigilator in groups of about 15. In order to complete all tests, each subject required six sessions of approximately 45 minutes duration each. The order of presentation of tests varied randomly between groups.

RESULTS

Factor Analysis

The correlations between the 23 variables were subjected to a principle components analysis. There were six principle components with latent roots greater than unity, and these were rotated for the orthogonal solution to the varimax criterion. The six factors together accounted for 65 per cent of the total variance. The variables which had loadings greater than .30 on the first three factors are listed in Table 1.

TABLE 1

ORTHOGONAL FACTORS ROTATED TO VARIMAX CRITERION

Variable	Factor I	Factor II	Factor III
Spelling: Word type			
Frequent-Regular	.867		
Infrequent-Regular	.867		
Frequent-Irregular	.902		
Infrequent-Irregular	.885		
Spelling: Error type			
Phonetic omissions	.789		
Phonetic additions	.648		
Phonetic transpositions		—.303	
Phonetic substitutions	.760		
Non-phon omissions	.767		—.352
Non-phon additions	.640	.420	
Non-phon transpositions	.320		—.304
Non-phon substitutions	.725	.325	
Incomplete		—.698	—.549
Unrecognisable	.674		
Memory			
Visual sequential			
Auditory sequential	—.480		
Intelligence			
Progressive matrices	—.555		
Reading			
ACER Speed	—.631		—.323
ACER Accuracy	—.789		
PAT Comprehension	—.804		
PAT Vocabulary	—.867		
Speed			
Coding		.359	
Writing Speed		.495	
Per cent of Variance	43.0%	7.3%	5.1%

Factor I, which accounted for 43 per cent of the total variance, was characterised by moderate to high loadings with respect to 12 of the 14 spelling variables and all four reading variables. The Progressive Matrices and auditory sequential memory also had moderate loadings on this factor, which may be described as a *verbal-intellectual* factor. Visual sequential memory was not related to this factor or to any of the other five factors. It is worthy of note that all phonetic and non-phonetic spelling errors, with the exception of phonetic transpositions, loaded significantly on Factor I. The product-moment correlation between total phonetic errors and total non-phonetic errors was .93.

Factor II, which accounted for 7 per cent of the total variance, was distinguished by few incomplete spellings, fast writing speed, high coding performance, the commission of certain types of spelling errors (non-phonetic additions and substitutions) and an absence of phonetic transpositions. This factor was described as *writing speed.*

Factor III, which accounted for 5 per cent of the total variance, was also distinguished by few incomplete spellings, and by low reading speed and an absence of certain types of spelling errors (non-phonetic omissions and transpositions). This may, perhaps, be described as a *carefulness* factor.

It was not possible to identify the remaining three factors, which together accounted for only 10 per cent of the total variance.

Analysis of Word Type

A rank order of difficulty was made of the 50 words in the spelling list, based on the number of correct spellings per word. The frequency of the 50 words was also ranked according to the frequencies quoted in the Thorndike-Lorge (1968) list. A Spearman rank order correlation was computed for the two rankings, which yielded *rho*= .80, significant at the .01 level.

A correlated t-test was calculated to test the difference between error scores on regular and irregular words. The mean error score for regular words was 10.66 (SD=6.20), and for irregular words, 12.33 (SD=6.28). The difference between these scores was not significant (t=0.32, df=92).

DISCUSSION

In the Reid and Hughes (1974) study, it was found that spelling achievement had a relatively high loading on a factor described as 'rote memory for verbal material'. The two memory variables which loaded significantly on this factor involved the reproduction of sequences of letters, which is a task very similar to the spelling of nonsense words. If memory is involved in these tests, therefore, it would seem to be specific to sequences of letters. As none of the other memory variables (associative memory, sentence completion, sentence recall, consequences) loaded significantly on this factor, it is difficult to justify Reid and Hughes' (1974) conclusion that spelling appears to involve 'a single common ability—that of short-term memory'.

The results of the present study indicate that if spelling involves a single common ability, it is not short-term memory, but rather a general verbal-intellectual ability. The only memory task which loaded on the same factor as spelling was auditory sequential memory, and its loading was not as high as most of the other variables. Visual sequential memory did not load on this factor, nor on any other factor.

The finding that phonetic and non-phonetic spelling errors intercorrelated very highly and loaded on the same verbal-intellectual factor indicates that these two types of error do not constitute independent categories. There does not seem to be any diagnostic value, therefore, in classifying students according to the phonetic accuracy of their spelling errors, as suggested by Simon and Simon (1973).

Day and Wedell's (1972) finding that visual and auditory memory were not associated either with phonetic or non-phonetic spelling errors was supported by the results of the present study. The findings of Nelson and Warrington (1974) did not, however, receive any support; reading, spelling, phonetic errors and non-phonetic errors were all associated with the same general factor. The only type of spelling error which appeared to be related to different abilities was the incomplete category.

The relation of word-type to spelling performance provided clear-cut results; for boys aged 13 to 14 years, accuracy of spelling was a function not of regularity of orthography, as defined by Lee (1960), but of word frequency. These results are consistent with those of Brown (1970) who used different criteria to ascertain frequency and regularity, an older population, and a different statistical analysis. The consistency of this finding adds weight to the already persuasive arguments against the usefulness of spelling reform (Chomsky, 1970; Gillooly, 1973).

The main implication of this study is that spelling instruction in the secondary school should be concerned more with infrequent words than with grapheme-phoneme correspondences. Since spelling is predominantly a verbal-intellectual task, it would seem that infrequent words would be easier to spell if they were made more familiar and more meaningful. It is suggested, therefore, that spelling instruction should form an integral part of vocabulary training. The importance of morphemic elements in English orthography (Chomsky, 1970) should not be ignored.

REFERENCES

ACER Silent Reading Tests (1963). Melbourne: Australian Council for Educational Research.

BANNATYNE, A. D., and WICHIARAJOTE, P. (1969). Relationship between written spelling, motor functioning and sequencing skills. *J. Learning Disabilities,* 2, 4-16.

BLOOMER, R. H. (1961). Concepts of meaning and the reading and spelling difficulty of words. *J. educ. Res.,* 54,178-182.

BLOOMER, R. H. (1964). Some formulae for predicting spelling difficulty. *J. educ. Res.,* 57, 395-401.

BROWN, H. D. (1970). Categories of spelling difficulty in speakers of English as a first and second language. *J. verb. Learn. verb. Behav.,* 9, 232-236.

CHOMSKY, C. (1970). Reading, writing, and phonology. *Harvard educ. Rev.,* 40, 287-309.

CRAMER, R. L. (1970). An investigation of first grade spelling achievement. *Elem. English,* 47, 230-237.

DAY, J. B., and WEDELL, K. (1972). Visual and auditory memory in spelling: an exploratory study. *Br. J. educ. Psychol.,* 42, 33-39.

GILLOOLY, W. B. (1973). The influence of writing—system characteristics on learning to read. *Reading Res. Q.,* 8, 167-199.

HOLMES, J. A. (1954). A substrata analysis of spelling ability for elements of auditory images. *J. exp. Educ.,* 22, 329-347.

LEE, W. R. (1960). *Spelling Irregularity and Reading Difficulty in English.* Slough: National Foundation for Educational Research.

LESTER, M. (1964). Graphemic-phonemic correspondences as the basis for teaching spelling. *Elem. English,* 41, 748-752.

LYLE, J. G. (1969). Reading retardation and reversal tendency: a factorial study. *Child Dev.,* 40, 833-843.

LYLE, J. G., and GOYEN, J. D. (1969). Performance of retarded readers on the WISC and educational tests. *J. abnorm. Psychol.,* 74, 105-112.

NAIDOO, S. (1972). *Specific Dyslexia.* London: Pitman.

NELSON, H. E., and WARRINGTON, E. K. (1974). Developmental spelling retardation and its relation to other cognitive abilities. *Br. J. Psychol.,* 65, 265-274.

PETERS, M. L. (1967). The influence of reading methods on spelling.*Br. J. educ. Psychol.,* 37, 47-53.

Progressive Achievement Tests. (1969). Wellington: New Zealand Council for Educational Research.

RAVEN, J. C. (1938). *Standard Progressive Matrices.* London: H. K. Lewis.

REID, N. A., and HUGHES, D. C. (1974). A factorial analysis of the P.A.T.: Reading Comprehension, Reading Vocabulary and Listening Comprehension Tests. *New Zealand J. Educ. Studies,* 9, 18-30.

RUDISILL, M. (1957). Interrelations of functional phonic knowledge, reading, spelling, and mental age. *Elem. School J.,* 57, 264-267.

SIMON, D. P., and SIMON, H. A. (1973). Alternative uses of phonemic information in spelling. *Rev. educ. Res.,* 43, 115-137.

SPACHE, G. (1940). Validity and reliability of the proposed classification of spelling errors. *J. educ. Psychol.,* 31, 204-214.

THORNDIKE, E. L., and LORGE, I. (1968). *The Teacher's Word Book of 30,000 Words.* New York: Teachers College, Columbia University.

TOWNSEND, A. (1947). An investigation of certain relationships of spelling with reading and academic attitude. *J. educ. Res.,* 40. 465-471.

(Manuscript received February 8, 1977)

Br. J. educ. Psychol., 47, 274-284, 1977

THE PROCESS OF TYPIFICATION IN CLASSROOM INTERACTION: MODELS AND METHODS

By D. H. HARGREAVES
(Department of Education, University of Manchester)

SUMMARY. Recent work on the process of typification and person perception in teacher-pupil relations is subjected to critical scrutiny. The similarities between several models, based on different methods and disciplines, are examined and an amalgamation of the models is proposed. It is argued that current models exhibit a number of important defects, especially in relation to contextual variations in typification and the changes in typification over time. A number of theoretical and methodological suggestions are made to improve models of classroom typification and interaction for both psychologists and sociologists.

One interesting feature in the historical development of the social and behavioural sciences is the substitution of a new concept for an older concept when both appear to refer to what is essentially the same phenomenon. A good example in social psychology is the replacement of 'gregarious instinct' by 'affiliative motivation'. To the cynical eye such conceptual transformations, especially where they arise on the boundaries between disciplines, may be seen as putting old wine into new bottles. On this view the changes indicate a concern with fashionable novelty or the imperialistic tendencies of competing disciplines as opposed to genuine conceptual innovation or advance. Whilst this may be a justified complaint in some cases, it must not blind us to what may be real changes in the assumptions, models and methodology which lie behind conceptual modifications. It is the object of this paper to examine one such conceptual change and to show how it relates to developments (or otherwise) in the models and methods by which we analyse the phenomenon.

In the last ten years psychologists, social psychologists and sociologists of education have all converged on the classroom as an important arena for educational research (Jackson, 1968; Rosenthal and Jacobson, 1968; Dreeben, 1968; Smith and Geoffrey, 1968; Kounin, 1970; Westbury and Bellack, 1971; Nash. 1973; Good and Brophy, 1973; Brophy and Good, 1974; Chanan and Delamont, 1975; Hargreaves, *et al,* 1975; Sharp and Green, 1975; Martin, 1976; Stubbs and Delamont, 1976; Nash, 1976). Much of this literature employs a range of well established concepts to analyse classroom life, such as attitude, self-image, role expectation, as well as newer concepts such as self-fulfilling prophecy and hidden curriculum. Indeed, the very popularity of such concepts rests upon their elasticity by which they be assigned various meanings and methods of measurement. One common theme throughout this literature is what teachers and pupils expect of one another and how they perceive one another. This phenomenon can be studied with the aid of the concepts of attitude or role at the interpersonal level and self-image (self-perception and role conception) at the individual level. Both levels may be involved in the one analysis, as in studies of the self-fulfilling prophecy.

More recently, and largely under the influence of interactionist and phenomenological social psychology and sociology, the concept of *typification* (including self-typification) has been introduced into the field. This concept, drawn from the work of Alfred Schutz (1932, 1973), encapsulates the traditional problem of the ways in which one person perceives, categorises, construes or types another person, and is closely allied to concepts such as attitude, role and self. It is generally accepted that this phenomenon (or perhaps set of phenomena) is of central significance in the analysis of classroom life and teacher-pupil relations. The issue at stake is the extent to which this is merely a conceptual novelty or a genuine conceptual advance which offers an improved model of classroom relationships as well as improved methods of investigation. The resolution of this issue involves a careful examination of contemporary writing on this topic.

One most influential, and relatively early, study of teacher-pupil relations in the classroom is that of Howard Becker (1952). He argued that teachers, like members of any service organisation, have an image of the 'ideal' client against which actual clients are matched and evaluated as 'good' or 'bad'. In the classroom, the good pupil is one who approximates closely to the teacher's ideal image, so presenting him with few problems. A bad pupil, by contrast, deviates markedly from the ideal, so reducing the teacher's experience of reward and success in the exercise of his profession. Becker postulates three elements in this model; the problems of teaching, discipline and moral acceptability. Good pupils conform closely to the ideal in displaying ability and successful learning, in conforming to classroom rules and regulations, and in not offending the teacher's moral standards. As befits an interactionist sociologist, Becker uses interview methods and his analysis concentrates on the way in which pupils from different social classes are judged against this ideal and differentially typified. The basic model is clear; typification is a matter of *ideal-matching.*

It is perhaps the rapid development of role theory at this period which accounts for the relative indifference shown by many investigators to the potentialities of Becker's model. In Britain, the first major study of role in an educational setting is that of Musgrove and Taylor (1969). The first part of this study(Taylor, 1962) is concerned with pupil expectations of teachers, but uses a different model and method. In place of Becker's interview, Taylor favours a content analysis of pupils' essays from which are then constructed scales in the format of an attitude test. There is some link with Becker's work in that each scale item begins with 'A good teacher . . .' which shows that both researchers are concerned with the pupil's ideal teacher. But for Taylor this ideal is held to constitute a set of role expectations. There is perhaps an implicit implication that pupils match actual teachers against this ideal, as Becker's teachers did with pupils, but this aspect is not developed. Indeed, the way in which pupils respond to actual teachers is totally neglected. The role scale diverges from Becker in that the good teacher is constituted as a bundle of characteristics, which are grouped together under the headings of teaching, discipline, personality and organisation. In this sense Taylor's study is akin to Hallworth's (1962) psychological investigation of teachers' rating of pupils' personalities, in which adjective check lists were subjected to factor analysis. This work was given some sociological sophistication by McIntyre *et al* (1966) who showed that the teachers' perception varied according to the social class intake of the school, thus forging a direct link with Becker.

This group of studies shares the same implicit model that teachers and pupils typify one another in terms of a set of characteristics. Actual teachers and pupils are typified as a unique configuration of such characteristics, and these typifications are constructed in the form of an identi-kit. This *characteristics model* appears at times to overlap with the ideal-matching model, but the relationships between the two are never explicated.

Ten years later the impact of symbolic interactionism and phenomenology was being felt by social psychologists in Britain. The basic tenets of interactionism indicated certain preferences for model and method. With regard to model, interactionists emphasise dynamic or process models which take into account changes in perception and typification over time and reflect contextual or situational variations. With regard to method, interactionists are committed to interview (like Becker) and participant-observation. In the early studies this preference for the qualitative is handled somewhat nervously, and so supportive quantitative data are usually collected simultaneously. Nash (1973) certainly showed his commitment to observation, but supported his observations of how teachers typify pupils with data derived from Kelly's repertory grid test, which became fashionable at this period and which has a strong congruence with interactionist tenets. Here, it appeared, was a technique which met interactionist canons in that it elicited the member's own constructs, rather than imposing them as happens in most attitude and personality tests, yet it also had the advantage of objectivity and quantifiability. (In their 'discovery' and appropriation of Kelly, sociological researchers almost entirely ignore the advances made in the field of person perception by experimental psychologists (eg, Warr and Knapper, 1968) and attribution theorists (eg, Jones and Davis, 1965); the convergence between the 'typification' approaches of sociologists and 'person perception' of psychologists continues to be surprisingly superficial). In the Kelly model as used by Nash, teachers and pupils typify one another in terms of bi-polar dimensions. According to Nash, primary school teachers in his study typified pupils in terms of three common constructs:—*hardworking—lazy, mature—immature, well behaved—poorly behaved,* though individual teachers employ additional idiosyncratic constructs. In a later study (1976b) Nash asserts that pupils typify teachers in terms of six major constructs: *keeps order—unable to keep order, teaches you—doesn't teach you, explains—doesn't explain, interesting—boring, fair—unfair, friendly—unfriendly.*

Clearly Nash's use of the Kelly model is closely allied to earlier work using the ideal-matching and characteristics models. But do the Nash studies mark any real advance on earlier work? Certainly Nash adds to the content of our knowledge, but I want to argue that this important contribution is nevertheless confirmatory rather than innovative. This is so because his theoretical position is only superficially interactionist in orientation and involves no change of model, and because his methodological position gives priority not to observation but to the Kelly technique, which is a variation on older methods.

Whilst there are differences between the Kelly and characteristics models, notably in the emphasis given by Kelly to the member's own constructs, it is the similarities between the two models that are striking. Both see the typification process as an imputation of a set of attributes. Once committed to the Kelly technique and model, Nash uses his observation to support and illustrate how the constructs are used in the daily interactions of teachers and pupils. Observational data do not in any way surpass or extend the data derived from the repertory grid test; they are used to corroborate not to go beyond the Kelly model. There is no indication of the interactionist's commitment to observation as a means of generating more adequate models (Glaser and Strauss, 1967). In consequence, his model of typification is not a dynamic one, which explores both how teacher and pupil typifications change over time and vary according to the situation.

This argument can be amplified by showing that the Kelly method, which

Nash seeks to incorporate into an interactionist analysis, is in fact subject to the same objections that the interactionist would bring against the 'psychometric' methods used to support the characteristics model. First, Nash assumes that the constructs provided by the teacher which, on common sense grounds, sound similar are in fact essentially the same. For instance, he tells us (1973, p22) that *hard working—lazy* and *tries hard—slacks* are the same, because 'their communality was obvious'. This is a very similar assumption to that embodied in the attitude or personality test, that the respondent will interpret a particular term or item in the same way. But it it justified to assume that 'hardworking' is synonymous with 'tries hard' and that 'lazy' is synonymous with 'slacks'? Perhaps so, but this is the single illustration that Nash gives of his interpretive work in this study; we are asked to take the rest on trust and have no means of checking whether his assumptions of sameness constitute a distortion of the data. The work of one of my own students (Stead, 1976) suggests that we must be cautious in accepting such assumptions. Stead shows that the construct 'slow', which has an apparently 'obvious' meaning is used by teachers in at least two different senses. Sometimes it is used to mean 'dull' as in the phrase 'slow learner'; but on other occasions the term is used to refer only to speed of working, as in the phrase 'he writes painfully slowly', without any reference to intellectual capacity. Again, the terms 'bright' and 'intelligent' appear to have the same meaning in Nash's work, but Stead shows that there is a distinct difference in meaning, 'intelligent' sometimes bearing a meaning 'high IQ' but 'bright' having the additional reference to the pupil's manner as 'alert' or 'perky'. How can we be sure that Nash has interpreted the meanings of the constructs correctly?

Second, Nash assumes, like psychometricians, that the meanings are not only constant but also transpersonal, that is, shared by all teachers. Many teachers often use the same term, such as 'likeable', but what evidence do we have that they see a common set of qualities as constituting 'likeability'?

Third, Nash assumes that each construct can, without distortion, be regarded as a simple continuum in bi-polar dimensional form. Among the six constructs he claims are used by pupils to typify teachers are *keeps order—unable to keep order,* and *friendly—unfriendly.* My own experience with pupils suggests that these are not two extreme positions on a simple continuum along which any teacher can be located. If that were so, where for instance would one locate the common definition of a teacher by pupils as being excessively concerned to keep order ('too strict with us') or as being more friendly than is desired ('familiar')? Moreover a host of typifications, such as 'grumpy', 'sarcastic', 'serious', 'nasty' and 'cold' are assumed to lie at one point ('unfriendly') on the same, single continuum.

Fourth, Nash assumes that the triadic elicitation technique of the Kelly repertory grid test is an unproblematic method for eliciting the teachers' or pupils' constructs. In asking the teacher how one of three pupils differs from the other two, Nash believes it is unimportant to know what context is supplied by the teacher in framing an answer to the question. To respond to the tester's request the teacher must 'fill in' or supply some implicit context or situation in which the pupils can be differentiated. One presumes that teachers supply some kind of classroom context, but they are free to supply other contexts, such as the playground or assembly, or the dinner-table. Yet the way in which teachers typify in such contexts probably varies considerably. The child who is typified as a 'chatterbox' or 'noisy' in the classroom may not be typified in this way in a playground setting. And even in the classroom setting, the pupil who is typified as 'inattentive' in a situation where the teacher is lecturing may not be so typified during project work. Nash has no control over the contexts implicity supplied by his respondents; he merely assumes that they supply the same basic context, or that the context supplied makes no difference. The second alternative is the more damaging for it makes the massive assumption that all significant typifications are necessarily *transcontextual,* and that the test can tap the constructs on a context-free basis.

In contrast to this, the interactionist is highly sensitive to
(a) situational or contextual variations in the typification process and
(b) the changing nature of typifications as they develop over time.
On the first point it is an axiom of Mead's interactionist social psychology that the meaning of an object, person or event is a function of a person's intention towards it. The meaning of a piece of chalk varies according to whether I intend to use it on the blackboard (writing implement) or throw it at a recalcitrant pupil (missile). Teacher and pupil typifications of one another vary, then, according to the changing intentions and interests of each in relation to the changing contexts of their interaction. Intentions and interests vary not only by locale (classroom, playground, assembly) but also within the same locale (the constantly changing activities within the 'same' lesson). And the data-gathering process (the interview, the repertory grid test, etc) must be viewed as yet another context with its own structure and dynamic. The profound implications of this will be examined later. On the second point, the interactionist should be sensitive to the changing nature of typifications over longer time sequences. Typifications are not fixed entities which can be tapped arbitrarily at any point in time. Rather typifications have a developmental career which mirrors the career of the teacher-pupil relationship. Like relationships, typifications have histories. Nash's model does not take these matters sufficiently into account and must be seen as another version of the characteristics model.

Nash is not the only interactionist to present the characteristics model of typification in the guise of an interactionist model. Furlong's recent paper (1976) justifiably attacks earlier models of pupil conduct in the classroom for failing to take account of the constantly shifting nature of meanings in interaction and coins the useful concept of the 'interaction set' to cope with the problem he exposes. But then in the second half of the paper when he turns directly to pupil typifications of teachers, he ironically abandons this admirable sensitivity to situational variations by arguing that pupils typify teachers in terms of two bi-polar dimensions of *strict—soft* and *effective—ineffective,* thus reviving the characteristics model. We learn nothing at all of the particular contexts in which pupils do, or do not, employ such typifications. Similarly Woods (1976) takes an interactionist approach to the examination of pupils' views of school which seeks to give priority to pupils' interpretive processes, but on the matter of pupil typifications of teachers he presents the data in the form of simple dichotomies such as *helpful—unhelpful, provides variety—boring, allows more freedom—little freedom.* The same criticism can be brought against Rist (1970), who like Furlong and Woods uses observation and interview. All three change methodology, but not model.

Gannaway's (1976) study, however, changes both. On the basis of interview and observation, he constructs a dynamic model by proposing that teachers are progressively typified by pupils in a given sequence. The teachers are, in effect, subjected to a systematic series of tests by pupils, the first of which is the extent to which the teacher keeps order, followed by the 'Can we have a laugh?' and the 'Does he understand pupils?' tests. The model is sketched very tentatively by Gannaway, but it has considerable potential, for this model can be readily developed to take account of both situational variations and the longer term career of the pupils' typifications. This is the first British study to specify a dynamic model of pupil typifications, even though Werthman's pioneering study (1963) of such a model has been available for nearly fifteen years.

With regard to teacher typifications of pupils, Sharp and Green (1975) offer a model which draws directly on the phenomenological concepts of Alfred Schutz, namely his distinction between consociates and contemporaries. Their analysis offers some interesting insights into the way in which primary school teachers typify 'difficult' or 'deviant' pupils as against 'normal' or 'good' pupils, but the model is relatively unelaborated and certainly does not constitute a dynamic model which takes account of situational features or the long-term career of the typification process. My own recent work (Hargreaves *et al,* 1975) on teacher typifications, using interview methods, attempts to offer a model that includes the career of the typification process in that it claims that the process passes through three analytically distinct stages: first, a stage of speculation in which the teacher hypothesises that the pupil is a certain kind of person; a second stage of elaboration, in which the early hypothesis is put to the test and then extended and elaborated; and third, a stage of stabilisation in which the teacher develops a highly complex but stable typification which is then resistant to change. A fourth stage, that of type transformation, is posited but not developed. This model certainly captures the dynamic career of the typification process, but it has two major deficiencies. Its principal application is to the 'deviant' pupil and it remains an open question whether it can profitably be applied to the 'normal pupil' and it fails to take account of situational variations in any adequate way. It seems reasonable to conclude that whereas an interactionist social psychology has certainly initiated a change of methodology in the investigation of the typification process, there has been relatively little development of new models. Much interactionist work offers no change of model at all; other work points forward to more dynamic models, but no general, adequate model is yet available.

It may be, however, that the different models discussed in this paper—the ideal-matching model, the characteristics model, and the dynamic interactionist model—are not really in competition at all. Although they arise from different assumptions and methodological preferences, each model perhaps exposes one facet of the typification process to the exclusion of other aspects. In other words, can the different models be combined into a more coherent whole? I tentatively suggest a combined model as follows.

(1) In the (relatively unusual) situation in which one meets another type of person for the first time (eg, a pupil's very first encounter with a teacher), that person will be typified by matching against the first person's *stereotype* of that type of person (stereotype-matching).

(2) Where one is familiar from previous experience with a type of person, then during early encounters that person will be typified by a process of matching against *typical* others. For example, a teacher meeting a class of first-year pupils will type them in relation to his common sense understanding, based on previous experience, of what constitutes typical first-year pupils (typical-matching).

(3) Where several persons of a given type are currently encountered, each individual will be typified also by a process of matching against these peers. For instance, a pupil will typify a teacher by comparing him with other

teachers known to him, and a teacher will typify a pupil by comparing him with other pupils (peer-matching).

The process in 1-3 will be essentially speculative or provisional. At this stage the person is not grasped very fully as a unique individual (though that he really is so is presumed), and is seen in terms of a configuration of characteristics, many of which are fairly crude bi-polar dimensions. It is at this stage that the characteristics model represents the typification state fairly accurately. The processes underlying this state are outlined in 1-3 above.

(4) As a person becomes experienced in dealing with typical others, he forms a conception of an ideal other. This will arise because typifications are evaluative not merely descriptive. So pupils will come to develop a conception of an ideal teacher against which actual teachers can be compared. This ideal need not be clearly or sharply defined. Ideal-matching then becomes an integral feature of typifying.

(5) After a process of elaboration, the typification of an individual other with whom we interact frequently becomes stabilised. Normally at this stage the typification becomes extremely complex and contains much more than a constellation of characteristics, and is not well represented by the characteristics model. The characteristics are moulded together into a relatively coherent whole, but are also set within a much broader framework. A teacher typifying a pupil at this stage understands the characteristics he attributes to the pupil within a framework of his understanding of the pupil's motives, the biography he imputes to the pupil, his knowledge of the pupil's home background, his relations with peers and so on. A characteristics model hints at these elements on occasion, as when the characteristic 'popular' draws upon the teacher's understanding of the pupil's peer relations; but the model abstracts these characteristics arbitrarily and artificially from the teacher's much fuller formulation of the pupil as a person.

(6) This 'stabilised typification' must be treated as an analyst's construction. We must beware of appearing to claim that in some way it exists in the typifier's head as a coherent whole. Rather, it is the analyst's aggregation of the enormous variations in the typifications employed by the typifier in the variety of contexts in which he actually typifies the other. Each context leads the typifier to typify in given ways as a function of the structure of that context and the intentions of the persons who interact in that context. The whole stabilised typification could never in principle be applied in any *one* context, and it is for this reason that the stabilised typification must be seen as an artificial construction of the analyst. It is the contextual variations in typifying that have received least research attention and which in consequence are not adequately integrated into our current models.

This argument suggests that the characteristics model most accurately represents typification at the relatively early stages of acquaintanceship and that it is seriously distortive of the nature of typification both in more intimate relationships and in contextual applications. The reason for this distortion in the model may lie in the methodological base of the data on which the model is built. The characteristics model, whether it springs from 'psychometric' sources (attitudes or personality test, the repertory grid test) or 'interactionist' sources (interviews, observation), usually rests upon data which are collected in a situation where teacher or pupil expresses his typifications of the other in what I shall call a *third party relation.* By this I mean that both test and interview require the respondent to typify the other outside the direct face-to-face relationship between teacher and pupil. The respondent typifies the other who is not present to a third party, the researcher. Yet the model constructed on these data is then expected to throw light on and to replicate the process of typification as it occurs in direct interaction in the classroom.

My argument is that it is a questionable assumption that data collected in one kind of situation (third party talk) can be used to represent the typification process that takes place in a very different kind of situation (face-to-face encounter). However, it may much more adequately represent the typification process as it occurs in other situations which are structurally similar to the interview/test situation. In fact, third party talk is not unique to research settings; there are 'natural' third party situations in the lives of teachers and pupils. Pupils frequently talk to one another about teachers, as do teachers about pupils. Playground gossip and staffroom gossip are moderately well reflected by the characteristics model derived from research-respondent talk, because all three are examples of third party talk. But the characteristics model and natural third party talk stand in a highly problematic relation to the typification process as it occurs in classroom interaction. The relation between what I have called third party talk and direct face-to-face interaction is in fact systematically analysed by Schutz (1932, 1973), though paradoxically none of the writers who have used his notion of typification and its associated concepts, such as Sharp and Green (1975) or Hargreaves *et al* (1975), appears to have recognised the methodological implications of this. The difference is essentially between what Schutz calls 'consociates' and 'contemporaries'. The term consociate is restricted to the other person with whom one is in a direct face-to-face relationship. Schutz describes this as a 'we-relation', because the two persons are conjoined in a special way in direct interaction. Each treats the other as 'thou' rather than a 'he' and so the we-relation can be described as a reciprocal thou-orientation. In such a relation each party has an immediate grasp of the other as a conscious being, whose conscious life becomes accessible through a range of symptoms, such as his words, his facial expression and his gestures. And each party knows that the other grasps him in the same way. Since each tunes into the other, their experience forms a common course or stream. They are 'bound in genuine simultaneity'. Contemporaries, on the other hand, are persons who are not bodily present, and one's relation with an absent person is not at all like that with a consociate. One is not bound up with a contemporary, who becomes a 'he' not a 'thou'. The we-relation is replaced with a he-relation.

Schutz then argues that the typification process is different in these two types of social relation. Typifications as a set of characteristics certainly play an important role in a relation between consociates, but the paramount feature of this relation is that the other is grasped as a unique person whose conscious being manifests itself vividly before one's eyes. Typifications are a help in grasping the other as a unique person whose situated conscious life and experience unfold before one, but they are constantly subordinated to, and modified by, the vivid reality of the other's presence. Contemporaries, by contrast, are grasped *predominantly* through the typifications. The contemporary must be grasped reflectively, because he is no longer present, with the effect that the other is treated as an object not as a subject, as appertains in the we-relationship. Who the contemporary is must be reconstructed out of the typifications which remain as a sedimentation of past experience, and the precise form of that construction will rest upon one's motives for trying to recall or reflect upon the absent other. The significance of this for my analysis is obvious. Natural third party talk (such as staff gossip about pupils) and a respondent's third party talk (whether in the form of a test or an interview) both tap those aspects of the typification process by which a person conceives the other as a contemporary. As researchers we cannot then create a model based on the typification of contemporaries and then transpose it to another situation as a model of consociate relations. Yet this is precisely what our current models of teacher-pupil typifications do.

Schutz's contrast between the consociate and the contemporary is, in application to the classroom, too simple to fit the facts. This is shown in the following list of common situations:

(a) Teacher and pupil have a private face-to-face interaction (Schutz's 'pure' consociate relation);

(b) The teacher interacts briefly with one pupil but in the presence and hearing of other pupils;

(c) The teacher addresses the class as a whole, glancing only briefly at individual pupils;

(d) The teacher observes a pupil who is not aware that he is under observation;

(e) The teacher talks to some colleagues about a pupil (natural third party talk);

(f) The teacher writes a report on a pupil who is not physically present (Schutz's 'pure' contemporary relation).

Clearly, between the two 'pure' types stands a range of common and important types of social relation in which typification plays an important part.

Most of our methods for gathering typification data have the structure of an attitude test; an interview can easily become an interpersonal, oral attitude test. Typification research may throw little light on how people interact, just as much attitude research shows a notoriously low correlation with people's conduct (Deutscher, 1965), and for the same reason, the testing situation is a special situation of its own which has too few equivalents in the everyday life of people. There is an overlap between the two settings, even though they cannot be taken as near-identical, and the importance of Schutz's analysis is that it shows that typifications derived from third party talk (where the other is a contemporary) do play some part in direct interaction (where the other is a consociate). But there are additional factors in operation in direct interaction, which are not at work in third party talk, so the major research task before us now is to specify and analyse those additional features and their relation to the typification process. Then, perhaps, we could develop what the interactionists have claimed they seek, namely a model which will allow us to understand contextual variations in the typification process. To achieve this we must follow Schutz's lead and strive for a model of interaction, not just a model of typification. As long as we persist in attempting to separate the latter from the former we are bound to err. And it is perhaps in a model of classroom interaction that psychologists and sociologists can find their meeting point.

As a possible first step we might refine all those methods which rely on third party talk, for this is a hallowed tradition which contains forms as diverse, in some respects, as the attitude test, the repertory grid test and the interview. Great technical ingenuity may be required to develop new research methods which will allow us to get closer to the we-relation of direct interaction between teachers and pupils in the classroom. But our traditional methods are also capable of considerable refinements that are more adequate to the task. Observational data, which have hitherto been used largely for illustrative purposes, can be used to refine and modify models that originally derive from third party talk. We can adapt other methods by attempting to *contextualise* them. Nash's use of the Kelly repertory grid test, as I pointed out earlier, asks decontextualised questions in which the respondent must typify the other in an abstract and situationally dislocated way. Yet there is no reason why contexts cannot be systematically built into the triadic elicitation test. For example, one might ask a teacher how he differentiates one of three pupils from the other two in a particular setting, such as the classroom, the gymnasium, assembly, the playground, at the dinner table and so on. A systematic comparison of similarities and differences in the typification process in each setting would then yield considerable insight into situational variations. Similarly, the test could be used to examine the typifications that arise at different points in time in the one setting, such as in different subjects, at the beginning and end of lessons, in activities that have different pedagogical structures (eg, when the teacher is giving a demonstration, conducting a question-and-answer session, supervising group work, etc). Likewise, in an interview with a pupil instead of asking him a series of general decontextualised questions about what he thinks of his teachers, we can ask what he thought of his teachers at particular places (in the classroom, in the corridor, in the playground, etc), and at particular points of time in the same locale (when the teacher praised his work, asked him a question, reprimanded him, etc). Such data could then be compared and contrasted with more natural third party talk, such as a non-directive discussion among pupils about their teachers.

Once in possession of a range of such data, we could critically examine, rather than merely make assumptions about, the relation between aspects of the typification process in different contexts, and elaborate the important but neglected distinction made by psychologists (Warr and Knapper, 1968) between 'episodic' (or context-bound) typifications and 'dispositional' (or transcontextual) attributes.

In contextualising the typification process in this way we must capitalise on the interactionists' strength and be careful to avoid imposing the researcher's definition of the significant contexts and events—a practice common in psychological work—for these may not be the most salient ones for teachers and pupils. If we are to grasp the dynamic changes in the contextualised typification process we must ensure that we capture the participants' definition of the most significant events and their meanings for understanding these events. Nash (1973) sought to contextualise his Kelly-based teacher typifications of pupils by supplementing the data with his observiations on lessons in progress; but it is Nash himself, not the teacher or pupil, who decides which events are the significant ones and the highly evaluative terms in which they are described. The analysis is inadequate in that this does not throw sufficient light on how or in what terms the participants typify in relation to events they regard as salient. The relation between Nash's account and that of the members remains problematic, because Nash does not ask for the participants' version, which may differ from his own.

Schutz's model of typification and interaction is of an 'armchair' variety, not based on research. If we can accept that his analysis carries a powerful implicit critique of our current models of typification and interaction in classrooms, and can respond by adapting our methods, whether 'psychometric' or 'interactionist', then we can create new models covering relations between both contemporaries and consociates, situational variations, and the long-term career of typifications. Such a model would be a marked advance on any of those currently available and would be a more adequate representation of the complex phenomena of life in classrooms.

ACKNOWLEDGMENTS—The author gratefully acknowledges the help of Norma Swift, Stephen Waterhouse and Peter Woods who made critical comments on an earlier draft.

REFERENCES

BECKER, H. (1952). Social class variations in the teacher-pupil relationship. *J. educ. Sociology,* 25, 451-65.

BROPHY, J. E., and GOOD, T. L. (1974). *Teacher-Student Relationships.* New York: Holt, Rinehart and Winston.

CHANAN, G., and DELAMONT, S. (1975). *Frontiers of Classroom Research.* Windsor, England: N.F.E.R.

DEUTSCHER, I. (1965). Words and deeds: social science and social policy. *Soc. Prob.,* 13, 232-54.

DREEBEN, R. (1968). *On What is Learned in School.* Reading, Mass: Addison-Wesley.

FURLONG, V. (1976). Interaction sets in the classroom: towards a study of pupil knowledge. In STUBBS, M., and DELAMONT, S. (Eds.), *Explorations in Classroom Observation.* London: Wiley.

GANNAWAY, H. (1976). Making sense of school. In STUBBS, M., and DELAMONT, S. (Eds.), *Explorations in Classroom Observation.* London: Wiley.

GLASER, B. G., and STRAUSS, A. L. (1967). *The Discovery of Grounded Theory.* London: Weidenfeld and Nicholson.

GOOD, T. L., and BROPHY, J. E. (1973). *Looking in Classrooms.* New York: Harper and Row.

HALLWORTH, H. J. (1962). A teacher's perception of his pupils. *Educ. Rev.,* 14, 124-33.

HARGREAVES, D. H., HESTER, S. K., and MELLOR, F. J. (1975). *Deviance in Classrooms.* London: Routledge and Kegan Paul.

JACKSON, P. W. (1968). *Life in Classrooms.* New York: Holt, Rinehart and Winston.

JONES, E. E., and DAVIS, K. E. (1965). From acts to dispositions: the attribution process in personal perception. In BERKOWITZ, L. (Ed.), *Advances in Experimental Social Psychology,* Vol. 2. New York: Academic Press.

KOUNIN, J. S. (1970). *Discipline and Group Management in Classrooms.* New York: Holt, Rinehart and Winston.

McINTYRE, D., MORRISON, A., and SUTHERLAND, J. (1966). Social and educational variables relating to teachers' assessments of primary school pupils. *Br. J. educ. Psychol,* 36, 272-279.

MARTIN, B. W. (1976). *The Negotiated Order of the School.* Toronto: Macmillan.

MUSGROVE, F., and TAYLOR, P. E. (1969). *Society and the Teacher's Role.* London: Routledge and Kegan Paul.

NASH, R. (1973). *Classrooms Observed.* London: Routledge and Kegan Paul.

NASH, R. (1976a). Pupils' expectations of their teachers. In STUBBS, M., and DELAMONT, S. (Eds.), *Explorations in Classroom Observation.* London: Wiley.

NASH, R. (1976b). *Teacher Expectations and Pupil Learning.* London: Routledge and Kegan Paul.

RIST, R. (1970). Student social class and teacher expectations. *Harv. educ. Rev.,* 40, 411-451.

ROSENTHAL, R., and JACOBSON, L. (1968). *Pygmalion in the Classroom.* New York: Holt, Rinehart and Winston.

SCHUTZ, A. (1932). *The Phenomenology of the Social World.* London: Heinemann.

SCHUTZ, A. (1973). *The Structures of the Life-World.* London: Heinemann.

SHARP, R., and GREEN, A. (1975). *Education and Social Control.* London: Routledge and Kegan Paul.

SMITH, L. M., and GEOFFREY, W. (1968). *Complexities of an Urban Classroom.* New York: Holt, Rinehart and Winston.

STEAD, J. (1976). Consensuality amongst teachers in their typification of pupils. Unpublished M.Ed. dissertation, University of Manchester.

STUBBS, M., and DELAMONT, S. (Eds.), (1976). *Explorations in Classroom Observation.* London: Wiley.

TAYLOR, P. H. (1962). Children's evaluations of the characteristics of the good teacher. *Br. J. educ. Psychol.,* 32, 258-66.

WARR, P. B., and KNAPPER, C. (1968). *The Perception of People and Events.* London: Wiley.

WERTHMAN, C. (1963). Delinquents in school: a test for the legitimacy of authority. *Berkeley J. Sociol.,* 8, 38-60.

WESTBURY, I., and BELLACK, A. A. (1971). *Research into Classroom Processes.* Columbia: Teachers College Press.

WOODS, P. (1976). Pupils' views of school. *Educ. Rev.,* 28, 126-37.

(Manuscript received November 26, 1976)

Br. J. educ. Psychol., 47, 285-295, 1977

LEXICAL CHARACTERISTICS OF ADOLESCENT PROSE: WRITER AND TASK EFFECTS

By E. MARGARET MOORE
(School of Psychology, Ulster College)

SUMMARY. Eight hundred and nineteen essays were obtained from 14-year-old and 17-year-old pupils of five Northern Ireland grammar schools. Lexical analyses of the scripts are reported here. Measures were calculated for word productivity; lexical diversity; voice, modal and aspectual characteristics of finite verbs, and for various pronominal and negator usages. Subgroup differences were analysed for pairs of subsamples discriminable by discourse mode (narrative-descriptive *vs* expository-argumentative) or by the writers' age, sex, Cattell Exvia or Anxiety scores.

INTRODUCTION

Despite a recent revival of interest in written compositions of adolescents among British psycholinguists and educationists (Department of Education and Science, 1975; Britton *et al*, 1975), quantitative data on their microstructures and correlates are scanty. Detailed information on developmental trends in prose writing abilities and on their interactions with cognitive and affective processes would seem essential for an adequate pedagogy. The present paper reports some findings for lexical measures derived from essays written by adolescent pupils in the third and sixth years of five co-educational Northern Ireland grammar schools (Moore, 1976). The linguistic indices chosen for investigation were thought to be of both stylistic and psychological interest, and to be sensitive both to intellectual and affective writer characteristics and task demands. Several hypotheses were advanced concerning the effects on these indices of task type (narrative-descriptive *versus* expository-argumentative) and the writer variables of age and sex.

It was hypothesised, following Moffett (1968) and Rosen (1969), that the narrative-descriptive mode would, by reason of its temporospatial syntactic structure and relatively slight lexical restriction, prove generally easier than the argumentative-expository. It was therefore predicted that pupils would tend to write more and exhibit greater lexical diversity on the former task. Since action is the essence of narrative, higher proportions of active to passive verbs were hypothesised for that mode, in line with the data for published prose reported by Svartik (1966). Similar considerations underlay a prediction of a higher proportion of dynamic to stative verbs in narrative writing. It was also assumed that the findings for adult samples of higher ratios of personal pronouns to finite verbs (Svartik, pp52-57) would hold for adolescent narratives as compared with expositions.

It was considered that sixth-formers would be enabled by their larger and more highly organised vocabularies to write longer essays and achieve higher lexical diversity scores in both discourse modes. On the other hand, it seemed probable that increased mastery over passivisation and impersonal constructions (resulting in an age-related decrease in proportions of personal pronouns to finite verbs) might be reflected only in exposition, which might be seen by older writers as requiring an indirect style more strongly than an 'exciting' story (Gibson, 1966). Similarly, it was considered that lexical changes assumed to reflect increasing capacity for speculative thought would be most clearly marked in expository-argumentative writing. Extrapolation from the work of Inhelder and Piaget (1958) and Peel (1971) suggested the likelihood of increased ratios of verbs expressive of possibility rather than fact to all finite verbs in older pupils' expositions. Increased ratios of all negators and of negators of 'verbs of possibility' in particular were also predicted for sixth year expositions.

There is substantial evidence for female superiority in word fluency on narrative writing tasks in childhood and early adolescence (La Brant, 1933; Harrell, 1957; Percival, 1965). In view of sex-linked differences on other perseverative tasks (Broverman *et al,* 1968), it was assumed that adolescent girls would tend to write more than male co-evals in both discourse modes. Brown and Gilman (1966) suggested that the use of first-person grammatical subjects and negators in argumentative writing might be indicative of aggression. It was hypothesised that boys would be likelier to express such aggression than girls and would score higher on such indices in the expository-argumentative composition. On the other hand, it was assumed that the independence and assertiveness commonly associated with the male (Kagan and Moss, 1962; Hutt, 1972) would be reflected in narrative-descriptive writing by a stronger preference for active verbs in general and dynamic active verbs in particular, in clauses with first-person grammatical subjects.

METHOD

Sample

Co-educational grammar schools in Northern Ireland outside the Greater Belfast area were approached in the spring of 1971. Schools known to have intakes of fewer than four forms at the time of the investigation were excluded. It was hoped to obtain data from six schools, but in the event it was possible to secure data from only five. It was intended to use complete year-groups, but in practice some fee-paying third-year pupils (none of whom had passed the 11-plus selection tests) were excluded in three schools for varying reasons, and the science lower sixth was withheld in a fourth.

Size of school intake proved a poor predictor of size of sixth year, as there were considerable variations in growth rate and numbers leaving after O-levels. Numbers in subgroups homogeneous in respect of school, age and sex, ranged from 9 to 76, and totalled 819, after the exclusion of pupils who failed to complete both research tasks required of them and those who wrote fewer than 200 words on either composition. Mean ages for these subgroups at the time of writing ranged from 14 years 2 months to 14 years 6 months for the third-formers, and from 17 years 3 months to 17 years 9 months for the sixth-formers. It was decided not to use up a research session for the collection of up-to-date cognitive test data. Eleven-plus verbal reasoning quotients, which were obtained for 368 pupils in four of the five schools, showed subgroup means ranging from 111.32 to 125.00. These figures probably exaggerate the abilities of the adolescents sampled, since they exclude fee-payers who had not been entered for the selection procedure because they had been considered unlikely to succeed. However, the samples were probably somewhat above average in verbal reasoning abilities. There was no reason to suspect systematic differences in such abilities between sex or age groups.

The High School Personality Questionnaire (1963, Form A) was administered to pupils of three schools, the anglicised version being as yet unavailable at the time of this study. Subgroup means (after corrections for sex and age) ranged from 4.7 to 6.9 for Exvia and from 4.6 to 6.9 for Anxiety.

Procedure

Pupils in schools 1 and 5 performed both writing tasks. Pupils in schools 2 and 4 wrote compositions elicited by stimulus A (see Appendix), which was intended to encourage argumentative-expository writing, and subsequently completed HSPQ forms. Pupils in school 3 wrote compositions elicited by stimulus B, designed to stimulate narrative-descriptive writing and based on a passage in Chapter 28 of Paul Bowles' novel, *The Sheltering Sky* (1949), and likewise completed HSPQ forms. The stimulus passages were made available to subjects in duplicated form as well as being read aloud by the investigator.

Linguistic Indices

(1) *Number of words.* Hyphenation was standardised with reference to *Chambers Twentieth Century Dictionary, New Edition* (1972) before counting, and hyphenated words were counted as single units.

(2) *Lexical diversity index (pooled).* Like indices 4 *et seq,* this measure was calculated for a corpus of 200 words, comprising the first and last hundred words of each composition. Carroll's (1964, p64) statistic, which divides the total of different word-types by the square root of twice the total of all words (tokens), was employed. Before data input, words which would normally be distinguished by capitalisation (eg, 'proper' *Hill* and 'common' *hill*) or pronunciation (eg, past tense and present tense of *read)* were allotted distinctive entries on data sheets. Semantic differences were ignored in view of the brevity of the samples and the difficulty of obtaining objective criteria for differentiation.

(3) *Lexical diversity index (segmental mean).* This was the mean of indices calculated separately for the first and last hundred-word-segments of each composition.

(4) *Personal pronoun ratio.* Personal pronouns, including reflexives and possessive pronouns (eg, *mine)* but excluding possessive adjectives (eg, *my)* and impersonal usages of *it,* were totalled, and the sum was divided by the total number of finite verbs.

(5) *Overt I/my ratio.* The dividend for this statistic consisted of the total occurrences of the words *I* and *my,* regardless of grammatical function. The total number of finite verbs again served as divisor.

(6) *I/my — subject ratio.* A count was made of the instances when the overt or implicit grammatical subject of a finite clause was *I* or *my* plus a noun, and this total was divided by the total of finite verbs.

(7) *Active verb ratio.* The total number of active finite verbs was divided by the total of all finite verbs.

(8) *I/my active verb ratio.* The total of active finite verbs with overt or implicit *I* or *my* plus a noun as grammatical subject was divided by the total of all finite verbs having such a subject.

(9) *Other subject active verb ratio.* The total of active finite verbs with overt or implicit grammatical subjects other than *I* or *my* plus a noun was divided by the total of all finite verbs having such subjects.

(10) *Dynamic verb ratio.* The total of dynamic active finite verbs was divided by the total of active finite verbs.

(11) *I/my dynamic verb ratio.* The total of dynamic active verbs with overt or implicit *I* or *my* plus a noun as grammatical subject was divided by the total of all active finite verbs with such a subject.

(12) *Other subject dynamic verb ratio.* The total of dynamic active verbs with overt or implicit grammatical subjects other than *I* or *my* plus a noun was divided by the total of all active finite verbs with such subjects.

(13) *Reality verb ratio.* The frequencies of verbs expressing statements in the forms of fact (whether relating to the past, present or simple future) as distinct from statements of possibility, questions, wishes or commands, were

summed. The total was divided by the total of all finite verbs.

(14) *Negator verb ratio.* The words 'no', 'not', 'none', 'nothing', 'nobody', 'no-one', 'nowhere', 'never' were summed together with the affixes 'never-' and 'non-'. The total was divided by the total of all finite verbs.

(15) *Negator of possibility ratio.* Those negators (as defined in 14) which occurred in clauses which *before* negation could be considered expressive of possibility rather than reality (see 13) were summed. The total was divided by the total of all finite verbs.

Data analysis

The calculation of several indices was greatly assisted by the use of a computer program (Nar One) devised by Mr W. Blackwood of the Queen's University, Belfast, Computer Laboratory (after Veldman, 1967). This counted word frequencies in prose input and reordered types in quasi-alphabetical arrangement.

Variables 8 and 11 (I/my active verb ratio and I/my dynamic verb ratio) could not be computed for 67 and 68 scripts respectively, which yielded zero divisors. The missing cases were, however, scattered among subgroups and did not reduce the numbers of subgroup comparisons which could be made for task, age or sex differences. The criteria suggested by Quirk *et al,* (1972) were adopted in distinguishing dynamic from stative verbs.

The further subdivision of subgroups in schools 2, 3 and 4 into high and low scorers on HSPQ Exvia and Anxiety (those who scored above and at or below the overall sample medians of 5.6 and 6.1) resulted in one Low Exvia and one Low Anxiety cell which contained fewer than five pupils (two each in the case of variables 8 and 11). Such cells were excluded from statistical analyses. The Wilcoxon Signed Rank Test was used in the measurement of task effects and the Mann-Whitney in the measurement of age, sex and personality effects (Siegel, 1956).

RESULTS

The ranges of median scores obtained by subgroups on each of the 15 lexical indices under review are shown in Table 1.

Task effects

Outcomes of hypotheses and unpredicted trends are summarised in Tables 2 and 3. As with other variables, a 5 per cent level of significance (on one-tailed tests for differences in the predicted direction and two-tailed tests for unpredicted trends) has been accepted.

Age effects

Included in the lists of unpredicted trends for age, sex and personality traits are all cases in which *all* subgroup differences for one or both tasks were found to be in the same direction (regardless of the number of subgroup comparisons which could be made), and also those cases in which the proportion of subgroup differences in one direction reached the 5 per cent level of significance on the binomial test (Siegel, 1956). It is recognised that since subgroups were variously interrelated (some by sex, some by school, some by age, and some by none of these characteristics) the application of this criterion is purely arbitrary.

Tables 4 and 5 show the outcome of hypotheses relating to age differences and unpredicted trends.

TABLE 1
LINGUISTIC INDICES: MINIMUM AND MAXIMUM SUBGROUP MEDIANS FOR YEAR AND TASK SETS.

Index	Narratives		Expositions	
	3rd Yr.	6th Yr.	3rd Yr.	6th Yr.
1 No. of words	319.5 439.0	267.5 411.0	267.0 346.0	306.0 375.0
2 Lexical diversity index (pooled)	5.63 6.15	5.80 6.28	5.50 5.80	5.78 6.10
3 Lexical diversity index (segmental)	4.685 5.004	4.596 4.985	4.650 4.809	4.738 5.056
4 Personal pronoun ratio	0.678 0.828	0.717 0.782	0.238 0.404	0.208 0.326
5 Overt I/my ratio	0.621 0.692	0.626 0.701	0.115 0.152	0.084 0.163
6 I/my — subject ratio	0.517 0.613	0.469 0.621	0.089 0.167	0.047 0.149
7 Active verb ratio	0.959 1.000	0.958 1.000	0.764 0.870	0.789 0.867
8 I/my active verb ratio	0.951 1.000	0.955 1.000	1.000 1.000	1.000 1.000
9 Other subject active verb ratio	1.000 1.000	0.969 1.000	0.750 0.845	0.750 0.850
10 Dynamic verb ratio	0.524 0.666	0.500 0.643	0.219 0.380	0.286 0.429
11 I/my dynamic verb ratio	0.533 0.694	0.500 0.643	0.219 0.440	0.286 0.429
12 Other subject dyn. verb ratio	0.551 0.600	0.484 0.600	0.276 0.400	0.313 0.429
13 Reality verb ratio	0.684 0.789	0.645 0.756	0.348 0.466	0.390 0.600
14 Negator verb ratio	0.076 0.097	0.091 0.116	0.059 0.116	0.050 0.118
15 Negator poss. verb ratio	0.033 0.043	0.000 0.067	0.000 0.050	0.000 0.056

TABLE 2
TASK EFFECTS: PREDICTED RESULTS AND NUMBERS OF SUBGROUP PAIRS YIELDING SPECIFIC OUTCOMES

Variable	Prediction	Outcomes			Stat. sign. outcomes in predicted direction
		Exp.+	Narr.+	E=N	
1 No. of words	Narr. plus	2	6	0	4
2 LDI (pooled)	Narr. plus	1	7	0	4
3 LDI (segment.)	Narr. plus	4	3	1	1
4 Pers. pron. rat.	Narr. plus	0	8	0	8
7 Active verb rat.	Narr. plus	0	8	0	8
10 Dyn. verb rat.	Narr. plus	0	8	0	8

TABLE 3

TASK EFFECTS: UNPREDICTED TRENDS

Variable	Outcomes			Stat. sign. outcomes within trend
	Exp.+	Narr.+	E=N	
5 Overt I/my ratio	0	8	0	8
6 I/my subject ratio	0	8	0	8
8 I/my act. verb ratio	8	0	0	4
9 Other subj. act. verb rat.	0	8	0	8
11 I/my dynamic verb ratio	0	8	0	7
12 Other subj. dyn. verb ratio	0	8	0	8

TABLE 4

AGE EFFECTS: PREDICTED RESULTS AND NUMBERS OF SUBGROUP PAIRS YIELDING SPECIFIC OUTCOMES

Variable	Prediction	Outcomes			Stat. sign. outcomes in predicted direction
		6th+	3rd+	6th=3rd	
1 No. of words	6th plus	8	6	0	7
2 LDI (pooled)	6th plus	12	2	0	9
3 LDI (segment)	6th plus	12	2	0	8
4 Pers. pron. ratio	3rd plus (Exp.)	1	7	0	2
7 Active verb rat.	3rd plus (Exp.)	4	4	0	1
13 Reality verb rat.	3rd plus (Exp.)	6	2	0	0
14 Negator verb rat.	6th plus (Exp.)	5	3	0	0
15 Neg. poss. v. rat.	6th plus (Exp.)	6	2	0	0

TABLE 5

AGE EFFECTS: UNPREDICTED TRENDS

Variable	Outcomes			Stat. sign. outcomes within trend
	6th+	3rd+	6th=3rd	
12 Other subj. dyn. v. rat. (Narr.)	0	6	0	0
13 Reality verb rat. (Narr.)	0	6	0	0

Sex effects

The outcome of hypotheses relating to sex effects is summarised in Table 6 and related unpredicted trends are listed in Table 7.

TABLE 6

SEX EFFECTS: PREDICTED RESULTS AND NUMBERS OF SUBGROUP PAIRS YIELDING SPECIFIC OUTCOMES

Variable	Prediction	Outcomes			Stat. sign. outcomes in predicted direction
		B+	G+	B=G	
1 No. of words	G plus	1	13	0	6
6 I/my subject rat.	B plus (Exp.)	2	6	0	1
8 I/my act. v. rat.	B plus (Narr.)	1	5	0	0
11 I/my dyn. v. rat.	B plus (Narr.)	6	0	0	5
14 Negator ratio	B plus (Exp.)	1	7	0	0

TABLE 7

SEX EFFECTS: UNPREDICTED TRENDS

Variable	Outcomes			Stat. sign. outcomes within trend
	B+	G+	B=G	
4 Pers. pron. ratio	2	12	0	4
10 Dynamic verb ratio (Narr.)	6	0	0	2
11 I/my dyn. v. rat. (Exp.)	8	0	0	2
13 Reality verb rat. (Narr.)	6	0	0	1
15 Neg. poss. v. rat. (Narr.)	0	6	0	0

Personality effects

No main effects were predicted for HSPQ Anxiety or Exvia with regard to the lexical variables discussed here. Unpredicted trends discriminating High Exvia and Low Exvia groups are reported in Table 8. No such trends emerged for Anxiety.

TABLE 8

EXVIA EFFECTS: UNPREDICTED TRENDS

Variable	Outcomes			Stat. sign. outcomes within trend
	HiE+	LoE+	HiE=LoE	
1 No. of words (Narr.)	0	3	0	0
2 LDI (pooled) (Narr.)	3	0	0	0
8 I/my act. v. rat. (Narr.)	0	3	0	0
9 Other subj. act. v. rat. (Narr.)	3	0	0	0

DISCUSSION

Task differences

Parsimony requires that specific stimulus characteristics be considered in relation to obtained task differences before generalised discourse mode effects are inferred.

Paragraphal pre-structuring of the town-planning task may have reduced its apparent difficulty, thus inflating word-productivity and lexical diversity scores, which failed to show the predicted superiority on narratives. Moreover, the injunction to produce an 'exciting' story may have stimulated the use of word repetition for dramatic effect and thus differentially depressed the within-segment lexical diversity measure for narratives.

Again, the solitary situation of the narrator in the story-completion task may well have been responsible for task-related differences in several *I-my* measures, whilst his helpless situation probably invited the high proportion of rescue themes which raised the *I*-plus- passive constructions to a level possibly above-average for narrative.

Specific stimulus characteristics do not, however, seem to be responsible for the obtained overall superiority in proportions of personal pronouns, active verbs and dynamic verbs for narratives.

Age differences

Word productivity failed to show predicted age increases — perhaps because greater facility among older writers was matched by greater self-consciousness and longer planning periods.

The non-appearance of an age-related decrease in active verb ratios on expositions may have been associated with an increased readiness among sixth-formers to assume an active town-planning role, which could have masked increasing facility in the use of the passive construction. Support for this view seems to be provided by the unpredicted fall in the assignation of dynamic roles to persons other than the story-teller in the narrative. The failure of predictions regarding negators and negators of possibility suggested by existing evidence for their increasing prominence in logical thinking during adolescence (Peel, 1971; Ward, 1972) may likewise spring from the fact that mature writers were more concerned in their expositions with advancing solutions, rather than with elaborating difficulties or countering imagined arguments.

A consistent, though slight, trend for higher sixth form scores on 'Verbs of Possibility' in narratives, but not (as predicted) in expositions, suggests that involvement in one other possible world is a poorer index of mature imaginative thinking than evidence of ability to move appropriately in argumentation between the world of reality and a series of possible alternatives. Pupils who took the latter course punctuated their prescriptive and conditional statements in the Blue-Print task with supportive factual statements concerning real schemes and their known results. In narrative writing they found it necessary to move beyond factual description and narration of objects and events in the fictional world in which they had been placed, to illuminate the reality of this fictitious 'actual' by reference to alternative possibilities, past, present and future.

Sex differences

Substantial support was given for the prediction of greater word productivity among girls, although subgroup differences were often non-significant.

The efficacy of the dynamic/stative distinction as a sex discriminator is demonstrated not only with regard to narratives, but also (in the case of first-person subjects) to expositions—higher scores being consistently achieved by male groups. The psychological significance of the active-passive voice distinction is much less clear-cut. Given the higher score of girls on several pronominal measures, it seems likely that their preference for a direct, personal style (Warshay, 1969) may have inflated their first-person active verb ratios in narration. Failure of higher male scores to emerge in expositions for negation and first-person pronouns, as consideration of the Brown-Gilman study had suggested, probably reflects semantic differences between the present samples and the extracts from Thoreau and Emerson on which the former study was based. As has already been indicated, the element of argumentation was relatively small in the Blue-Print scripts, and negation was used as frequently to express obstacles to individual wish-fulfilment as contradiction of alternative views. The higher scores on 'verbs of possibility' and the subclass of 'negators of possibility' obtained by girls in narrative writing are, although unpredicted, intuitively compatible with the low self-confidence commonly attributed to females. Gleser *et al,* (1959) found a higher incidence of negators in the oral autobiographical narratives of women than of men.

Exvia differences

It is scarcely surprising that such personality-linked differences as emerged should occur in the less constrained and intellectually undemanding narrative task rather than in the exposition. Story-completion is, after all, a well-established projective technique. Since, however, these data rest on three subgroup comparisons from a single school, they must clearly be interpreted with caution.

Cattell and Klein (1975) note that 'the best markers for Exvia-Invia show casualness, low reaction to social inhibition, permissiveness, low rigidity and an unpunished optimism'. 'Casualness' seems consistent with the relatively low word productivity scores of pupils high in Exvia: 'low rigidity' is compatible with their superior lexical diversity scores which are also congruent with Di Scipio's (1971) findings for divergent thinking. The readiness of high scorers to write of the (usually helpful) activities of persons other than the story-teller in their narratives, and the avoidance by low scorers of themes of personal dependence (high *I-my* active verb ratios) together seem to reflect differentials in 'permissiveness' and 'unpunished optimism' in relation to interpersonal experiences.

Conclusion

The findings of this study would seem to suggest a need for more elaborately designed investigations capable of demonstrating the relative importance of various task and writer characteristics for performance of writing assignments commonly required of secondary school pupils.

ACKNOWLEDGMENTS—The teachers and pupils of the co-operating schools are thanked for their assistance.

REFERENCES

BOWLES, P. (1949). *The Sheltering Sky.* London: John Lehmann.

BRITTON, J., BURGESS, T., MARTIN, N., McLEOD, A., and ROSEN, H. (1975). *The Development of Writing Abilities (11-18).* London: Macmillan.

BROVERMAN, D., KLAIBER, E., KOBAYASHI, Y., and VOGEL, W. (1968). Roles of activation and inhibition in sex differences in cognitive abilities. *Psychol. Rev.,* 75, 23-50.

BROWN, R. W., and GILMAN, A. (1966). Personality and style in Concord. In SIMON, M. and PARSONS, T. H. (Eds.), *Transcendentalism and its Legacy.* Michigan: University Press.

CARROLL, J. B. (1964). *Language and Thought.* London: Prentice-Hall.

CATTELL, R. B., and CATTELL, M. D. L. (1969). *Handbook for the Jr.-Sr. High School Personality Questionnaire.* Champaign, Ill.: I.P.A.T.

CATTELL, R. B., and KLEIN, T. W. (1975). A check on hypothetical personality structures and their theoretical interpretation at 14-16 years in T-data. *Br. J. Psychol.*, 66, 131-151.

DEPARTMENT OF EDUCATION AND SCIENCE (1975). *A Language for Life.* (The Bullock Report). London: H.M.S.O.

DI SCIPIO, W. J. (1971). Divergent thinking: a complex function of interacting dimensions of extraversion-introversion and neuroticism-stability. *Br. J. Pschol.*, 62, 545-550.

GIBSON, W. (1966). *Tough, Sweet and Stuffy. An Essay on Modern American Prose Styles.* Bloomington and London: Indiana U.P.

GLESER, G. E., GOTTSCHALK, L. A., and WATKINS, J. (1959). The relationship of sex and intelligence to choice of words: a normative study of verbal behavior. *J. clin. Psychol.*, 15, 182-191.

HARRELL, L. E. (1957). A comparison of development of oral and written language in school age children. *Monogr. Soc. Res. Child Dev.*, 23, 3.

HUTT, C. (1972). *Males and Females.* Harmondsworth: Penguin.

INHELDER, B., and PIAGET, J. (1958). *The Growth of Logical Thinking from Childhood to Adolescence: an Essay on the Construction of Formal Operational Structures.* London: Routledge and Kegan Paul.

KAGAN, J., and MOSS, H. A. (1962). *Birth to Maturity, a Study in Psychological Development.* New York: Wiley.

LA BRANT, L. (1933). A study of language developments in children in grades 4 to 12 inclusive, *Genet. Psychol. Monogr.*, 14, 387-491.

MOFFETT, J. (1968). *Teaching the Universe of Discourse.* Boston: Houghton Mifflin.

MOORE, E. M. (1976). Some syntactic, lexical and phonetic characteristics of narrative and expository prose passages written by younger and older adolescent pupils in Northern Ireland grammar schools. Unpublished Ph.D. dissertation, Queen's University, Belfast.

PEEL, E. A. (1971). *The Nature of Adolescent Judgment.* London: Crosby Lockwood Staples.

PERCIVAL, E. A. (1975). The nature of ability in English composition. *Educ. Rev.*, 23, 189-204.

QUIRK, R., GREENBAUM, S., LEECH, G., and SVARTIK, J. (1972). *A Grammar of Contemporary English.* London: Longman.

ROSEN, H. (1969). An investigation of the effects of differentiated writing assignments on the performance in English composition of a selected group of 15/16-year-old pupils. Unpublished Ph.D. dissertation, University of London.

SIEGEL, S. (1956). *Nonparametric Statistics for the Behavioural Sciences.* New York: McGraw-Hill.

SVARTIK, J. (1966). *On Voice in the English Verb.* Janua Linguarum, Series Practica 53. The Hague: Mouton.

VELDMAN, D. J. (1967). *Fortran Programming for the Behavioral Sciences.* New York: Holt, Rinehart and Winston.

WARD, J. (1972). The saga of Butch and Slim. *Br. J. educ. Psychol.*, 42, 267-289.

WARSHAY, D. W. (1969). *Instrumental and Affective Language Styles in Sex Status.* Ph.D. dissertation, University of Idaho. Ann Arbor: University Microfilms.

(Manuscript received July 26, 1976)

APPENDIX

A *Expository-argumentative stimulus*

Please write these headings and the correct information on the first two lines on the first page of your answer booklet.

Your Name	Your School	
Your Form	Date of Birth	Today's Date

Do not write anything else till you are told.

Blue-Print for a new Town

Imagine that a planning committee has been set up to re-develop the town in which this school is situated. A large sum of money is available. You have been asked for your views on how it should be spent. Decide on FOUR major improvements to the town which should, in your opinion, have top priority. Explain clearly why you think they are important. Use words only—no drawings.

You will have half an hour to write your report. Try to be persuasive.

(I shall make a list of all the points you mention and bring a few copies of this list with me on my next visit to the school.)

Do not ask questions. If you can't spell a word, write what you think is best.

B *Narrative-descriptive stimulus*
(Preliminary instructions as for A above.)

Please read the short passage that follows, and go on with the story in your own words. Try to make your story as *exciting* as possible. This can be done by writing about exciting events or by writing about quite ordinary events in an exciting way.

You will have half an hour to write your story.

(I shall make a list of all your ideas for completing the story, with examples of bits of writing I thought were particularly exciting, and bring a few copies of this list with me on my next visit to the school.)

Do not ask questions. If you can't spell a word, write what you think is best.

"The blinding sea was there below, and it glistened in the silver morning light. I lay on the narrow shelf of rock, face down, head hanging over. My fingernails grated on the rock; I was certain I would fall unless I hung on with every muscle. But how long could I stay like that, suspended between sky and sea? . . . '

Br. J. educ. Psychol., 47, 296-304, 1977

A STUDY OF ATTITUDES TOWARDS MATHEMATICS IN RELATION TO SELECTED STUDENT CHARACTERISTICS

By R. F. KEMPA and J. M. McGOUGH*

(University of Keele, Staffordshire)

SUMMARY. Attitudes towards mathematics of over 300 first-year sixth-form students were studied in relation to the type of mathematics curriculum previously followed by them, to their choice of sixth-form courses and the students' intelligence, personality and mathematical achievement. The main findings are that no significant differences in liking for mathematics and in views about the difficulty of learning mathematics were found between students who had followed a 'modern' or 'traditional' mathematics curriculum, but that attitudes to mathematics were strongly correlated with students' mathematical bias as inferred from their choice of sixth-form subjects. As an additional aspect of the study, students' preferences for different modes of presenting mathematical information were examined, with three presentation modes being used: symbolic, graphical and verbal. Students' mathematical bias was found to be strongly associated with preference for the symbolic communication mode and anti-preference for the verbal mode.

INTRODUCTION

The period since 1960 has seen major changes in the nature of mathematics education in our secondary schools, resulting chiefly from the introduction of new mathematics curriculum projects, eg, the School Mathematics Project (SMP) and the Midlands Mathematical Experiment (MME), and the development by the GCE Boards of 'alternative' syllabuses both at O-level and A-level. It is thus not inappropriate to begin to evaluate the effect on the pupils of the introduction of the new mathematics curricula and syllabuses into the schools. To do so in relation to the cognitive outcomes of mathematical education has little meaning: courses and curricula vary extensively in their content and in the skills they try to foster, and hence a direct comparison of achievements cannot be made.

A better basis for an evaluation would be to consider attitudes to mathematics and mathematics learning as they appear in students having followed different curricula, but even this premise is not a straightforward one. It may be hypothesised that such attitudes vary not only with the mathematical experiences encountered by the student, but also with his choice of subjects, his mathematical achievement, his intelligence and other personal characteristics. If this is so, then the study of the relation between mathematics curricula and student attitudes is only one aspect of a wider-ranging enquiry into the relationship between mathematics attitudes and general student characteristics. This is the standpoint which has been taken in the design and execution of the present study.

---

*Now at the Department of Education, Manchester Polytechnic.

SCOPE OF THE INVESTIGATION

The study was carried out on pupils who had been taught according to the three different O-level mathematics syllabuses of the Joint Matriculation Board (1974); of these, Syllabuses A and B are essentially 'traditional' in character, although somewhat different in their mathematical content; Syllabus C, in contrast, is strongly biased to 'modern' mathematics. The confinement of the study to students from a single GCE Board gave the advantage of a direct comparability of their O-level grades in mathematics.

The attitude inventory used in this investigation was developed on the basis of the constructs and scales proposed by Husen in his 1967 *International Study of Achievement in Mathematics,* but contained several major modifications which are described elsewhere (McGough, 1975). It was thought important, when evaluating students' attitudes towards mathematics, to consider their 'mathematical bias' as indicated by the choice of different A-level subjects. Students reading arts subjects only, following their O-level mathematics courses, would be expected to have different attitudes to mathematics, compared with students opting for mathematics as a sixth-form subject. Such differentiation would be in line with Hudson's (1966) finding according to which students selecting mathematics-oriented A-level subjects display intellectual and personality characteristics different from those of other sixth-form students.

An additional aspect examined as part of the study was the extent to which students exhibited preferences for symbolic, verbal or graphical modes of presenting mathematical information. (This part of the investigation was based on an unpublished pilot study undertaken by K. S. Travers, R. W. Heath and L. S. Cohen in 1967, at Stanford University, California.) It was thought possible that this measure might differentiate between groups of students showing different mathematical biases and, possibly, attitudes. The tentative hypothesis was that a strong leaning towards mathematics would be associated with a preference for the symbolic representation mode and a disinclination to favour the verbal mode.

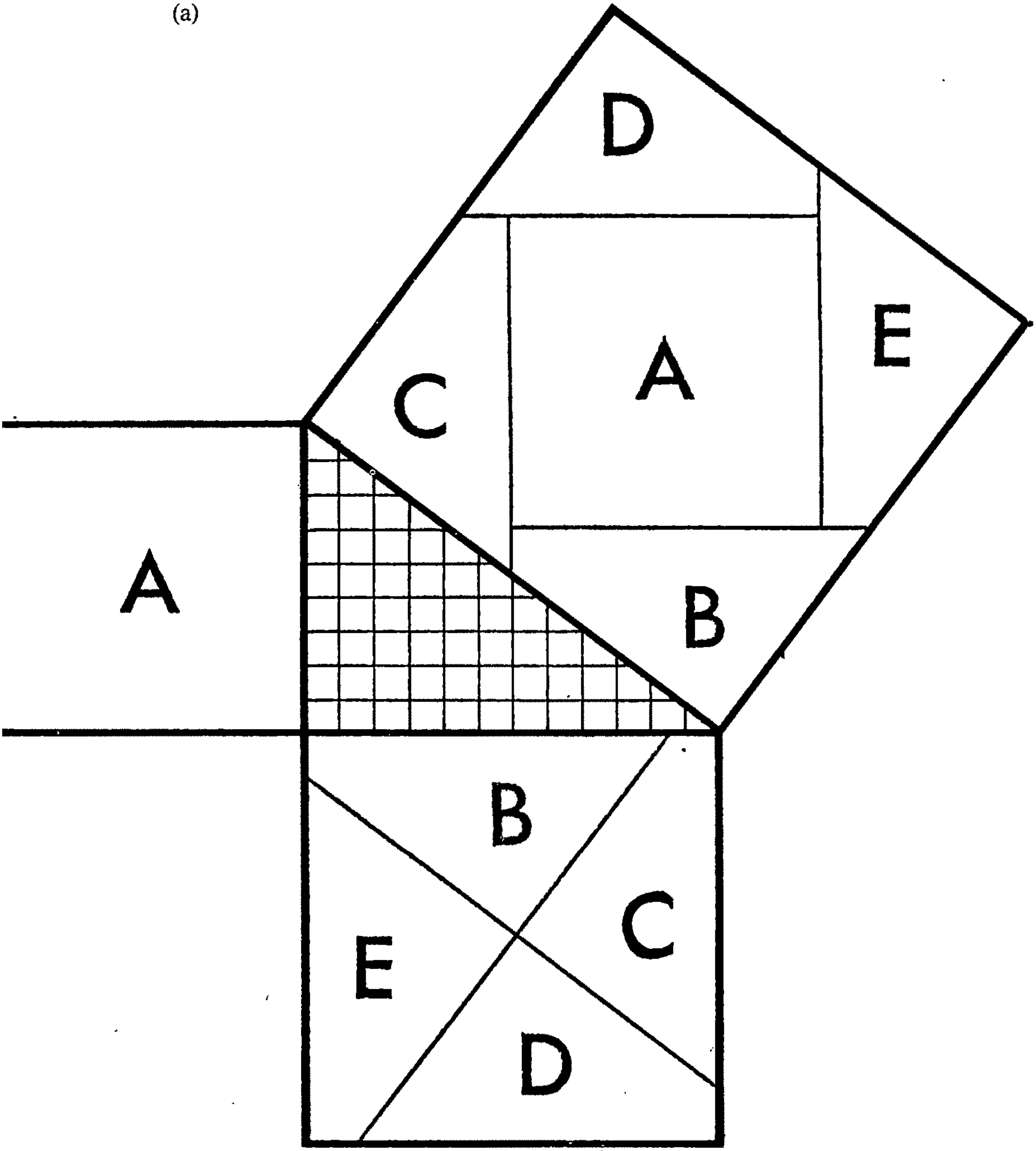
EXPERIMENTAL DETAILS

Instruments Used in the Study

The following instruments were developed: a test of 'cognitive preferences' in mathematics, a mathematics attitude inventory, and a student questionnaire seeking information about students' previous mathematical education, the range of subjects studied by them at the sixth-form level and the examinations for which they were studying. In addition, the AH4 test (Heim, 1970) and the Eysenck Personality Inventory (Eysenck, 1963) were administered. The latter two are standard instruments, and hence their underlying constructs are not described here.

The 'Cognitive Preference' Test: This was a test of 20 different items, each presenting a particular piece of mathematical information in a graphical, symbolic and verbal format. For example, the well-known theorem of Pythagoras could be expressed as follows:

(a)



(b) $(AC)^2 = (AB)^2 + (BC)^2$ where the triangle ABC is right-angled at B.

(c) In any right-angled triangle, the square of the length of the hypotenuse (longest side) is equal to the sum of the squares of the lengths of the other two sides.

All forms are essentially equivalent as far as their mathematical information is concerned, but differ in the manner in which this information is expressed.

Students, in responding to items of this type, were asked to express their relative preference for the different modes, allocating the rank scores 3, 2 and 1 to each of the alternatives given in a particular item. On the basis of the test as a whole, an overall preference score in relation to each of the information modes was obtained for each student.

In view of the ipsative nature of the cognitive preference test, ie, its property of leading to a set of preference scores the sum of which is constant, preference scores derived from it have no absolute significance. Moreover, none of the usual test evaluation procedures may be applied to this type of 'forced-choice' rating test. Ideally, an item forming part of the test should show no undue bias towards any of the three modes of expression; with respect to the population examined, this criterion was not fulfilled by the test. On the test as a whole, students gave the highest preference to the symbolic mode and the lowest preference to the graphical mode (the mean ratings were: 1.35 for graphical information, 1.98 for verbal information and 2.67 for symbolic information). Within each scale, the reliability of students' responses was high: split-half reliabilities were found to be 0.9 or higher. Despite the bias inherent in the test, as applied to the present population, each of the scales still discriminated adequately between different students. Therefore, results from the test were processed as part of the study.

The Attitude Inventory: The operational form of this comprised 33 Likert-type rating items based on the Husen (1967) inventory, and embraced the following construct sub-scales:

(i) Attitudes and views relating to the difficulty of learning mathematics (7 items; reliability 0.77);
(ii) Views on the importance and usefulness of mathematics in life (5 items; reliability 0.82);
(iii) Enjoyment of and liking for mathematical activities (12 items; reliability 0.93);
(iv) General attitudes to school work (9 items; reliability 0.80).

The scales were validated by the Guttman procedure and factor-analysis. Fuller details of the attitude inventory and other instruments are available (McGough, 1975).

The Sample

The sample consisted of 322 first-year sixth-formers from seven boys' grammar schools which, in the previous summer, had entered candidates for the JMB O-level mathematics examination (Syllabuses A, B and C).

For the part of the study which was concerned with differences between syllabus groups, the data from only six schools were used, giving a total of 291 students divided into three syllabus groups each of which contained students from two schools. The sizes of the groups were:
Syllabus A: N=117; Syllabus B: N=72; Syllabus C: N=102.

The 'mathematical bias' displayed by students in the choice of their sixth-form subjects was assessed on the following scale:

Bias level 1: Student is studying A-level arts subjects only.
Bias level 2: Student is studying at least one science subject but not mathematics.
Bias level 3: Student is taking A-level arts courses and a single mathematical subject.

Bias level 4: Student is studying science subjects, including single mathematics.
Bias level 5: Student is reading 'double' mathematics at A-level.

Administration

The administration of the tests was conducted by one of the authors during the summer term 1975, in the usual manner. Total testing time was approximately 100 minutes.

RESULTS AND DISCUSSION

The results of the study are presented and discussed in two separate sections dealing, respectively, with student attitudes to mathematics and with their cognitive preferences in relation to mathematical information.

Student Attitudes

Of immediate interest is the question of whether curricular differences, as encountered by students in their mathematical education before the sixth-form, are associated with different attitudes to mathematics. As a preliminary to this analysis, the distribution of students according to their mathematical bias (as defined above) and the type of mathematics course previously followed, was examined. This showed the three syllabus groups to be well matched in terms of mathematical biases. Hence, comparisons may be made directly between whole syllabus groups, without differentiation between bias levels.

TABLE 1

MEAN SCORES ON ATTITUDE SCALES FOR DIFFERENT SYLLABUS GROUPS

Attitude Variable		Total Sample (N=291)	O-level Syllabus A (Trad) (N=117)	B (Trad) (N=72)	C (Modern) (N=102)	F-Ratio P<
Difficulty/Ease of Learning Mathematics	Mean SD.	23.40 5.16	24.04 5.15	23.01 5.06	22.93 5.21	1.54 (NS)
Importance/Usefulness of Mathematics	Mean SD.	13.38 3.64	13.33 3.73	14.25 3.77	12.82 3.35	3.31 (0.05)
Enjoyment/Liking of Mathematics	Mean SD.	40.22 11.51	40.28 11.42	41.07 12.27	39.55 11.12	0.37 (NS)
General Liking of School/School Learning	Mean SD.	29.84 6.52	29.33 6.30	30.43 6.26	30.01 6.95	0.69 (NS)

Score ranges:				
	"Ease" scale:	7 to 35	"Enjoyment" scale:	12 to 60
	"Importance" scale:	5 to 25	"General Liking" scale:	9 to 45

The scale midpoints would represent "neutrality/indifference"; scores above the midpoints are 'positive' (in the 'difficulty/ease' scale they point to 'ease'); scores below the midpoint reflect a 'negative' orientation (or, for the first scale, 'mathematics is difficult').

Table 1 shows the mean scores on the four attitude scales, for the three syllabus groups. It is seen that no differences exist between the groups in relation to three of the measures. Although the fourth measure (importance of mathematics) does reveal a marginal difference between the groups, no real

TABLE 2

MEAN SCORES ON ATTITUDE SCALES FOR DIFFERENT MATHEMATICAL BIAS LEVELS

Attitude Variable		Total Sample (N=317)	Arts ← Bias Levels → Mathematics					F-Ratio P<
			1 (N=132)	2 (N=47)	3 (N=32)	4 (N=84)	5 (N=22)	
Difficulty/Ease of Learning Mathematics	Mean SD	23.34 5.17	22.57 5.71	24.06 4.31	23.03 4.58	24.21 4.78	23.59 5.46	1.62 (NS)
Importance/Usefulness of Mathematics	Mean SD	13.29 3.67	12.06 3.53	13.38 3.07	12.28 2.99	15.30 3.28	14.27 4.58	12.58 (0.001)
Enjoyment/Liking of Mathematics	Mean SD	39.67 11.70	31.47 9.12	37.00 9.02	45.75 10.27	47.88 6.91	54.41 5.19	72.47 (0.001)
General Liking of School/School Learning	Mean SD	29.68 6.40	29.24 6.48	30.92 6.00	28.25 6.92	30.10 5.96	30.09 7.49	1.10 (NS)

significance can be ascribed to this in view of the smallness of the student sample. The tentative conclusion must therefore be that curricular differences do not result in attitudinal differences.

The comparison of attitude ratings by students in different bias categories (Table 2) shows that two of the four scales distinguish significantly between the bias groups. These scales measure students' perception of the usefulness of mathematics, and their liking for and enjoyment of mathematical activities. The latter scale may, of course, in itself function as a crude measure of mathematical bias: it is unlikely that a student with little liking for mathematics would study mathematics or mathematics-related disciplines at the sixth-form level. Interestingly, the bias groups do not differ significantly in their perception of the difficulties of learning mathematics. It would appear that the 'difficulty' factor is far less important as a determinant of whether students opt for mathematical studies at the sixth-form level, than the 'enjoyment' factor.

The 'importance/usefulness of mathematics' scale produces an interesting distinction between 'scientists' (Bias groups 2 and 4) and 'non-scientists': the means for the former are distinctly higher than those for the adjacent bias groups. No doubt, this is a reflection of the science student seeing mathematics as a 'service subject'.

The 'general liking of school learning' measure fails to differentiate between the bias groups. This is not surprising inasmuch as all students in the sample had voluntarily stayed on at school after their O-level course. The absence of any correlation between the 'enjoyment of mathematics' measure and the 'general liking for school learning' one, even for students studying mathematics at sixth-form level, suggests that a person's liking for school learning depends on factors other than the mere liking of one particular academic subject.

The examination of attitude scores on the various scales in relation to IQ groupings, determined by the AH4 test, showed a significant pattern only in relation to 'liking/enjoyment of mathematics' variable: the higher the IQ grouping of the students, the more pronounced was their apparent enjoyment of mathematics learning. Personality variables were also found to be largely unrelated to attitude variables. The actual data are not presented here, but have been listed by McGough (1975).

TABLE 3

MEAN SCORES ON ATTITUDE SCALES FOR DIFFERENT ACHIEVEMENT LEVELS IN O-LEVEL MATHEMATICS

Attitude Variable		Total Sample (N=199)	O-Level Grades					F-Ratio P<
			A, B (N=31)	C (N=62)	D (N=36)	E (N=38)	Fail (N=32)	
Difficulty/Ease of Learning Mathematics	Mean	23.53	25.09	24.31	23.67	23.34	20.56	3.32
	SD	5.60	5.42	5.06	4.47	5.48	5.50	(0.02)
Importance/Usefulness of Mathematics	Mean	13.32	15.16	13.23	12.36	13.68	12.34	3.13
	SD	3.81	4.06	3.97	2.78	3.72	4.29	(0.02)
Enjoyment/Liking of Mathematics	Mean	39.84	49.16	44.53	37.08	36.34	28.99	23.18
	SD	11.56	7.96	9.24	8.82	10.31	10.46	(0.001)
General Liking of School Learning	Mean	30.03	30.94	31.26	28.89	28.11	30.34	1.93
	SD	6.39	6.14	5.44	7.26	5.88	8.05	(NS)

Table 3 presents the mean attitude scores for students in different mathematics achievement categories, defined in terms of O-level mathematics grades. Again, it is seen that only the 'enjoyment/liking of mathematics' scale produces a strong differentiation between the achievement categories: the higher the students' achievement in mathematics, the greater is their liking of mathematical activities. Differences on the 'difficulty' and 'importance' scales are slight only and, although marginally significant, are the result more of the pronounced views held by one or two achievement groups, than of steady trends. Students' general attitude to school learning appears invariant with mathematical achievement, probably for reasons already discussed.

Ideally, a two-way analysis of attitudinal characteristics with respect to mathematical bias and achievement would have been desirable, but could not be performed in view of the limited student sample and its uneven distribution. However, the provisional inspection of a two-way breakdown of the data did not suggest the existence of any major interaction effects.

Cognitive Preferences

Scores on the graphical, symbolic and verbal scales were not independent of each other in that the sum of any respondent's scores on the three scales was equal to a constant (120 in this case). In view of this, no correlational analyses could be performed, and only variations in scores on the three scales were examined. But even in this situation, the interdependence of the scale scores must be borne in mind.

The cognitive preference scores were examined in relation to the same variables used in the study of the attitude measures. The results are given in Tables 4 and 5. Additional data, not presented here, showed cognitive preference scores to be invariant with curricular differences and personality characteristics.

The patterns in the variation of cognitive preference scores with mathematical bias, are noteworthy (Table 4). The preference for the symbolic mode of expressing mathematical information is seen to increase steadily with increasing mathematical bias of the students. On the verbal communication scale, students in the lowest bias group exhibit, on average, a distinctly higher preference than the remaining students. Similarly, the two bias groups

TABLE 4

MEAN SCORES ON COGNITIVE PREFERENCE SCALES FOR DIFFERENT MATHEMATICAL BIAS LEVELS

Cognitive Preference Variable		Total Sample (N=315)	Arts ← Bias Levels → Mathematics					F-Ratio P<
			1 (N=131)	2 (N=47)	3 (N=32)	4 (N=83)	5 (N=22)	
Graphical/Diagrammatic Mode	Mean	26.89	28.05	27.38	25.34	25.78	25.32	3.14
	SD	5.86	6.08	6.19	5.89	5.21	4.78	(0.05)
Symbolic Mode	Mean	53.44	50.37	53.89	55.97	56.12	56.91	20.26
	SD	5.95	6.49	5.27	4.25	3.66	4.41	(0.001)
Verbal Mode	Mean	39.68	41.58	38.72	38.69	38.10	37.77	5.77
	SD	6.20	7.05	5.93	4.11	4.54	6.67	(0.001)

Score range on each of the cognitive preference scales: 20 to 60.

TABLE 5

MEAN SCORES ON COGNITIVE PREFERENCE SCALES FOR DIFFERENT ACHIEVEMENT LEVELS IN O-LEVEL MATHEMATICS

Cognitive Preference Variable		Total Sample (N=199)	O-Level Grades					F-Ratio P<
			A, B (N=31)	C (N=62)	D (N=36)	E (N=38)	Fail (N=32)	
Graphical/Diagrammatic Mode	Mean	26.63	26.29	26.27	26.11	27.74	26.94	0.55
	SD	5.67	5.71	5.61	5.39	5.62	5.27	(NS)
Symbolic Mode	Mean	53.60	56.16	54.10	55.08	51.95	50.46	5.89
	SD	5.78	4.02	5.28	4.12	7.33	5.92	(0.001)
Verbal Mode	Mean	39.76	37.55	39.63	38.81	40.31	42.59	3.27
	SD	6.02	3.74	5.32	6.06	7.20	6.39	(0.02)

comprising non-mathematics sixth-formers (Groups 1 and 2) show a greater leaning towards the graphical mode than the other groups, all of whom are studying mathematics at the sixth-form level.

Of particular interest is the observed relationship between the mathematical bias level and the pronounced tendency of students to prefer the symbolic communication mode. Without wishing to imply any cause-and-effect relationship here, the question presents itself of whether the cognitive preference leanings are in themselves an indicator of a person's predisposition to mathematics. It could be argued, for example, that a high preference for symbolic communication and, concomitantly, a relatively low preference for verbal communication modes are likely to lead a person to the choice of a discipline which, in its intellectual requirements, matches these psychological characteristics. This admittedly speculative point nevertheless merits attention.

This argument is reinforced by the data in Table 5 which analyses cognitive preference scores in relation to mathematical achievement at O-level. A fairly steady increase with mathematical achievement is noted for the preference scores on the symbolic communication scale, with the reverse trend applying to the verbal preference scores. It appears that the leaning to symbolic thinking modes which, it may safely be argued, is a concomitant of mathematical ability, expresses itself also in the preference for a symbol-based communication mode.

DISCUSSION

The study reported here was concerned with an examination of sixth-formers' attitudes to mathematics and mathematics learning and, in a preliminary way, with an investigation of their preferences for different modes of presenting mathematical information.

The main findings concerning attitudes to mathematics were as follows. Student populations who had followed a 'modern mathematics' curriculum did not exhibit attitudinal traits different from those who had taken more 'traditional mathematics' courses. If modern mathematics courses have as one intention to enhance students' enjoyment of mathematics, this intention was not achieved in relation to the present student sample which was, admittedly, rather limited.

Of the four attitude measures employed in this study, 'liking/enjoyment of mathematics' was found to be the one most strongly differentiating between student groups with different mathematical biases and those in different achievement categories. Interestingly, students' perception of the difficulties of mathematics learning was not found to be a correlate of their mathematical bias, and only weakly related to their previous mathematical achievement. This suggests that the liking of a subject may well be a stronger determinant of whether or not the subject is chosen for further study, than the learning difficulties associated with it.

Views about the usefulness and importance of mathematics seemed to differentiate more clearly between science and non-science sixth-former than relate in a straightforward way to students' mathematical bias.

The part of the study concerned with cognitive preferences for different modes of communicating mathematical information, was largely exploratory in terms of both the underlying concept and the instrument used. If follows from related studies in the field of science education (Heath, 1964; Kempa and Dube, 1973; Tamir, 1976). The present findings are that preference for the symbolic mode of presenting mathematical information is a significant correlate of students' mathematical bias and achievement. In turn, students with an arts bias tend to prefer the verbal communication mode. The area is open for considerable further work and may well yield information of value in the characterisation of 'disposition to mathematics'.

REFERENCES

EYSENCK, H. J., and EYSENCK, S. B. G. (1963). *Eysenck Personality Inventory. Personality Questionnaire.* London: University of London Press.

HEATH, R. W. (1964). Curriculum, cognition and educational measurement. *Educ. psych. Measur.*, 24, 239-253.

HEIM, A. W. (1970). *Manual for the AH4 Group Test of General Intelligence* (Revised Edition). Slough: N.F.E.R.

HUDSON, L. (1966). *Contrary Imaginations.* London: Methuen.

HUSEN, T. (Ed.) (1967). *International Study of Achievement in Mathematics.* London: Wiley.

JOINT MATRICULATION BOARD (1974). *Regulations and Syllabuses.* Manchester.

KEMPA, R. F., and DUBE, G. E. (1973). Cognitive preference orientations in students of chemistry. *Br. J. educ. Psychol.*, 43, 279-288.

McGOUGH, J. M. (1975). A study of attitudes towards mathematics in relation to curricular influences and selected student characteristics. M.Sc. Dissertation, University of Keele.

TAMIR, P. (1976). The relationship between achievement in biology and cognitive preference styles of high school students. *Br. J. educ. Psychol.*, 46, 57-67.

(Manuscript received November 18, 1976)

Br. J. educ. Psychol., 47, 305-311, 1977

LATERALITY AND INTELLECTUAL ABILITY: A JUST NOT NOTICEABLE DIFFERENCE?

By C. HARDYCK

(Institute of Human Learning, University of California, Berkeley)

SUMMARY. Previously reported work on laterality and intellectual abilities has suggested that the left-handed may be divided into two populations. When right- and left-handed are compared as groups on measures of intellectual performance, no differences are found. However, when children at the extreme lower end of the ability distributions are compared, a slight excess of left-handed has been reported. The present study selected, from a population of 7688 children, those at the lower end of IQ, scholastic aptitude, and reading ability distributions and found no difference in the frequency of handedness for this subgroup. Left-handed children are as likely to be at the upper end on an ability distribution as the lower end.

INTRODUCTION

The cumulative evidence of numerous studies assessing the relationships between handedness and intellectual and cognitive abilities strongly suggests that the image of the left-handed as cognitively or intellectually deficient is no longer given much credence. An overview of studies investigating the left-handed deficit hypothesis indicates (1) a remarkable lack of differences attributable to handedness, (2) a tendency on the part of many researchers to over-interpret miniscule differences based on small samples. Much of the most recent work hypothesising deficit in the left-handed (Levy, 1969; Nebes, 1971; Nebes and Briggs, 1974) has not been successfully replicated (Fagan-Dubin, 1974; Hardyck, 1976; Hardyck *et al,* 1976; Kutas *et al,* 1975).

Within this large body of evidence arguing that the null hypothesis is true, the reports on clinic samples, particularly on learning disability cases, remain as the last bastion of evidence for the association of left-handedness and deficit. From the early studies of Gordon (1921), Creak (1936), and Orton (1937) to the relatively recent studies of Burt (1958), Bryden (1970), Granjon-Galifret and de Ajuriaguerra (1951), Ingram (1959), and Wussler and Barclay (1970), studies investigating clinic samples have reported a slightly higher proportion of left-handed. This work has recently been examined by Annett and Turner (1974) who reviewed earlier work and reported on their findings in an attempt to resolve this discrepancy in reported results. Their hypothesis is an intriguing one: that the left-handed consist of two populations: one not discriminably different from the right-handed, and one deficient in intellectual and cognitive abilities.

As a test of this hypothesis, they examined 113 right-handed and 111 left-handed children between the ages of 5 and 11 on intelligence and reading measures. When the sample means were compared, no differences were found. When children at the extreme lower end of the ability distribution were classified as to handedness, a slightly larger proportion of left-handed were found. Annett and Turner argue that these individuals are from a population of left-handed which is developmentally delayed in the mastery of reading and verbal abilities.

The hypothesis is an interesting one and, if supported, would do much to clarify many of the apparent contradictions currently present within the research literature. Although it is difficult to assess the hypothesis with respect to learning disabled children since the samples used were drawn from community schools, the possibility that such individuals can be detected among children attending community schools is worthy of examination. It was with the hope of finding results similar to those reported by Annett and Turnei that we examined data on 7688 children tested on handedness, intellectual performance, and scholastic aptitude.

METHOD

Sample

The 7688 children were enrolled in grades 1 up to 6. The community is located on the California coast and is part of a larger urban area near two cities and has diverse business and industrial establishments as well as a large university. The population is diverse, with a large black and oriental population.

Measures

Handedness and Eyedness

Three measures of handedness and one measure of eyedness were taken: (1) *Handwriting.* The child was asked to write his name and the hand used to write was noted. (2) *Paper Cutting.* The child was asked to pick up a scissors and cut a card into two pieces. The hand used to operate the scissors was noted. (3) *Eyedness.* The child was asked to pick up a paper tube, the hand used to pick up the tube and the eye used for sighting through the tube were noted.

In general the children were remarkably consistent on the tasks, mixed handedness being so infrequent as to eliminate the necessity for any mixed-handed group.

Ability Tests:

(1) *Figure Copying Test.* The test is one of the least culture-affected tests available and successful performance on the task is known to be significantly related to readiness for the tasks of the primary grades. The test consists of 10 geometric forms, increasing in order of difficulty, which the child is asked to copy. There is no memory factor, since the test is before the child at all times. The test has a substantial correlation with intelligence tests, but is regarded as a measure of general cognitive development, as well as of perceptual motor ability.

(2) *Listening Attention Test.* In this test, the child is presented with an answer sheet containing 100 pairs of digits in sets of 10. He listens to a tape-recording presenting one digit every two seconds and is instructed to mark an X over the digit (of the pair) heard. This test is a measure of ability to pay attention, to keep to the appropriate place in a test and to make the appropriate response to what was heard from moment to moment.

(3) *Speed and Persistence Test.* This is a simple measure of test-taking motivation. It provides a measure of the subject's willingness to comply with instructions in a group testing situation and to mobilise effort. In the first part of the test, a child is asked to make Xs in a series of squares for 80 seconds. After a two-minute rest period the child is asked to better the first performance.

(4) *Lorge-Thorndike Intelligence Test.* This is a nationally standardised test of general intelligence and is acknowledged to be one of the better paper and pencil tests of general intelligence. Tests for grades 1-3 do not depend on reading, using pictorial items exclusively. The tests for grades 4-6 consist of two parts, verbal and non-verbal, scored separately. The following forms were used:

Grade 1:	Level 1, form B;

Grades 2-3: Level 2, form B;
Grades 4-6: Level 3, form B, verbal and non-verbal.

(5) *Scholastic Achievement.* Scholastic achievement was measured by a partial battery selected from the Stanford Achievement tests, consisting of the following sub-tests:

Grades 1, 2, 3: Word Meaning; Paragraph Meaning; Total Reading.
Grades 4, 5, 6: Word Meaning; Paragraph Meaning; Spelling; Work Study Skills; Language (Grammar); Arithmetic Computation; Arithmetic Concepts; Arithmetic Applications.

Data gathered on these children have been reported elsewhere (Hardyck, Goldman and Petrinovich, 1975; Hardyck, Petrinovich and Goldman, 1976) and will not be given here. It will suffice for the present analysis to state that no differences of any kind were found between the right- and left-handed. Out of all comparisons done, the largest total proportion of variance (over all measures) attributable to handedness was 2.1 per cent. The sample seems comparable to that of Annett and Turner in age and in the lack of differences between the means of the handedness groups.

To evaluate the hypothesis, the mean and standard deviation for all children were calculated for each grade, separately for each measure. The number of children with scores 1, 2 and 3 standard deviations (SD) below the mean were then separated on handedness. If these data are to show similar results to those found by Annett and Turner, we would expect a greater percentage of left-handed below the mean. Furthermore, the percentage of left-handed should increase as the distance from the mean increases, greater percentages of left-handed being present at the lower extremes of the distribution.

In addition to examining scores below the mean, we carried out a similar analysis for scores 1, 2, 2.25 and 2.50 standard deviations above the mean. The choice of these points was based on the observed distributions. There were no particular expectations for this analysis; it simply seemed worthwhile as long as scores in the lower extremes were being examined.

RESULTS

In Table 1, the frequencies of right- and left-handed for grades 1-6 is given. In contrast to Annett and Turner, who used similar size samples of right- and left-handed, the present comparisons were done on the frequencies of handedness occurring naturally in each grade. The comparisons reported here are by handedness groupings. Separate analyses done on groups of left-handed and left-eyed only and left-handed and right-eyed only did not differ from the handedness grouping. Since there seems little point in reporting subgroup analyses that do not differ from the total, only the handedness results are given here.

TABLE 1

THE DISTRIBUTION OF HANDEDNESS BY SCHOOL GRADE

	Grade					
Handedness	1	2	3	4	5	6
Right	1015	1120	1004	1020	875	1065
Left	125	116	104	108	92	116
Total	1140	1236	1108	1128	967	1181

In Table 2, the average percentage difference over all grades where measures are comparable is given for −1, −2, −3 and +2, +2.25, and +2.50 SD. In these comparisons, the frequency of right- and left-handed is each set equal to 100 per cent. The differences are then found by subtracting the percentage of left-handed beyond a given SD value from the percentage of right-handed found beyond that same value. These differences are then averaged over those grades where measures are comparable. A negative value in the table indicates that a greater percentage of left-handed are present beyond the SD value. In interpreting these differences it should be kept in mind that the total number of left-handed children is approximately one-ninth that of the right-handed.

TABLE 2

AVERAGE PERCENTAGE DIFFERENCES FOR MEASURES COMMON TO SCHOOL GRADES*†

Measure	Grades	−1 SD	−2 SD	−3 SD	+2 SD	+2.25 SD	+2.50 SD
Figure Copying	1-4	—2.23	—0.75		0.38	—0.23	0.34
Lorge-Thorndike IQ	1-3	—1.38	0.20				
Stanford Achievement Tests:							
Word Meaning	1-6	—1.33	—1.66				
Paragraph Meaning	1-6	0.15	0.13		—4.64	—3.71	
Total Reading Score	1-3	1.05			—3.12		
Lorge-Thorndike:							
Verbal IQ	4-6	1.84			—0.41		
Non-Verbal IQ	4-6	—3.21	—0.22		—0.86		
Stanford Achievement Tests:							
Spelling	4-6	—1.76	—2.44		1.96	2.00	
Language	4-6	—2.83	—1.49	—1.89	0.62	—0.40	
Arithmetic Computation	4-6	—2.01	—0.72		0.82	—0.13	
Arithmetic Concepts	4-6	0.97	—2.17		2.45	—0.13	
Arithmetic Applications	4-6	0.95	—0.40		0.60	—0.45	

*Differences are taken R — L. A negative value indicates a greater percentage of left-handed.

†Differences are cumulative. The percentage beyond −1 SD includes all individuals beyond −2 and −3 SD.

As can be seen by inspection of Table 2, a larger percentage for left-handed is more frequent in the −1 and −2 SD comparisons. For the −1 SD comparisons, the percentage of left-handed is larger 7 times and the percentage of right-handed 5 times. For the −2 SD comparisons, the percentage of left-handed is larger 8 times and right-handed 2 times. However, when the upper extremes of the distribution are examined, similar results are found. For a +2 SD comparison, the percentage of left-handed greater occurs 4 times and right-handed greater 6 times. For a +2.25 SD comparison, the percentage of left-handed is larger 6 times with that for right-handed larger occurring once. (Differences greater than −2 SD and +2.25 SD occurred only twice and seemed too small to warrant reporting.)

If these percentage differences are summed algebraically over all measures for the −1 SD, −2 SD, etc, comparisons, the result is even less supportive of the Annett and Turner hypothesis than the miniscule differences shown in Table 2. The algebraic summing and averaging indicates that a difference of less than 1 per cent is present for each comparison. The direction of the difference is uniformly that of a larger percentage present for the left-handed—in both the extreme lower and upper regions of the distribution. Given the differential frequency of occurrence of the right- and left-handed in our sample, the results indicate that one person in approximately half the grade levels is sufficient to account for the obtained differences.

The specific comparison Annett and Turner report in support of their hypothesis is the differential frequency of handedness found in children with a reading quotient 30 points or more below their Vocabulary IQ score. This comparison was not possible in the present study. For grades 1-3, there were no children with reading scores below −1 SD. In grades 4-6 there were also no children with Stanford Achievement word meaning scores below −1 SD. A reading ability-vocabulary discrepancy of this kind found by Annett and Turner is simply not present among our observations.

One additional analysis will be reported. The hypothesis that there will be a larger proportion of left-handed at the lower end of a distribution of measures implies that a mean difference between right- and left-handed will be present even though the difference may be quite small. Since the measures used in the present study were taken over several grades, each grade can be treated as a separate investigation of the hypothesis that the mean difference will be systematically lower for the left-handed. Given the relatively large size of our samples (the smallest N for any *t* test comparison of right- and left-handed is 967), the statistic

$$z = \frac{\Sigma t}{\sqrt{k}}$$

where *t* is the obtained value of the *t* statistic and *k* the number of independent experiments, can be evaluated using the normal curve (Winer, 1971). In Table 3, the number of independent determinations possible and the obtained *z* are given for each of the measures of the present study. As is evident in Table 3, no comparison even remotely approaches statistical significance.

TABLE 3

OBTAINED Z VALUES FOR ALL MEASURES TREATING EACH GRADE AS AN INDEPENDENT COMPARISON

Measure	Grades	Number of Independent Comparisons	z
Figure Copying	1-3	3	.8/3
Lorge-Thorndike:			
Total IQ	1-3	3	.729
Verbal IQ	4-6	3	.499
Non-verbal IQ	4-6	3	.173
Stanford Achievement Tests:			
Total Reading	1-3	3	.339
Word Meaning	1-6	6	.491
Paragraph Meaning	1-6	6	.296
Spelling	4-6	3	.051
Language	4-6	3	1.412
Arithmetic Computation	4-6	3	.713
Arithmetic Concepts	4-6	3	.392
Arithmetic Applications	4-6	3	.047

DISCUSSION

As was mentioned earlier, the hypothesis put forth by Annett and Turner was intriguing and the present study was undertaken in hopes of providing support for it. Unfortunately, the obtained results are not especially encouraging.

Comparisons on a much larger set of samples provides no support for the hypothesis that a group of left-handed may be identified at the lower extremes of a distribution of intellectual performance measures. Although the present study did indicate a slightly greater percentage of left-handed at the lower extremes of score distributions, the same result was found for the upper extremes of the score distributions. For both extremes, the percentage differences are extremely small.

Both the present samples and that of Annett and Turner were drawn from school systems serving normally functioning members of the school age population. Although there is no reason to suspect that the schools used by Annett and Turner differed in admissions criteria from our schools, the possibility cannot be eliminated. However, a more reasonable possibility is that the Annett and Turner results are primarily a function of sampling fluctuations. When the task faced is one of detecting what may be a relatively small difference, the possibility of misinterpreting sampling variations is quite high. As an illustration, the percentage differences for all measures for the fourth and fifth grade samples are shown in Table 4.

TABLE 4

RIGHT-LEFT PERCENTAGE DIFFERENCES FOR FOURTH AND FIFTH GRADE STUDENTS

Measure	Fourth Grade			Fifth Grade		
	−1 SD	−2 SD	−3 SD	−1 SD	−2 SD	−3 SD
Stanford Achievement Tests:						
Word Meaning	2.20	−0.73		−1.82	−3.86	
Paragraph Meaning	−0.44	−1.62		−2.98	−1.57	
Spelling	3.42	−3.67		−2.81	−0.91	
Language	3.66	−1.23	−2.57	−5.08	−2.46	−1.26
Arithmetic Computation	−1.81	−1.13		−3.78		
Arithmetic Concepts	8.24	2.44		−7.74	−2.88	
Arithmetic Applications	7.27	1.05		−3.93	−3.14	
Lorge-Thorndike:						
Verbal IQ	3.85			4.54		
Non-Verbal IQ	−0.07	−1.47		−4.51	1.18	

In looking for support for the hypothesis of a slightly greater frequency of left-handed at the lower extremes of a score distribution, quite different interpretations might have been generated if one of the above samples had been the only examination of the hypothesis. If only the fourth grade had been examined, the support for the hypothesis would have been indifferent at best, since the differences are small and change direction from measure to measure. If the fifth grade had been the only sample examined, confidence in the hypothesis would have been considerably higher.

The possibility remains that the Annett and Turner hypothesis is correct, but has not been adequately tested, since the determination of handedness by behavioural testing and the recording of familial history of handedness has not been reported for large samples of learning disability cases under conditions where the results are comparable to general population samples. The data of the present study are limited to children enrolled in community school systems, where no evidence can be found for the existence of this hypothesised subgroup. The possibility remains that handedness in some children may be related to a developmental delay of cognitive functioning. Such suggestions have been put forth by several investigators other than Annett (Satz and Sparrow, 1970) and the hypothesis is consistent with much of what is known about lateralisation of function and handedness. The actual evidence remains scanty and inconclusive. The Annett and Turner hypothesis remains an interesting one, but not one that can be supported by the present results.

ACKNOWLEDGMENTS—This study was supported by a grant from the Spencer Foundation. I would like to thank Barbara Cerny and Hilary Naylor for their help.

REFERENCES

ANNETT, M., and TURNER, A. (1974). Laterality and the growth of intellectual abilities. *Br. J. educ. Psychol.*, 44, 37-46.

BRYDEN, M. P. (1970). Laterality effects in dichotic listening: relation with handedness and reading ability. *Neuropsychologia*, 8, 443-450.

BURT, C. (1958). *The Backward Child* (4th edit.). London: University of London Press.

CREAK, M. (1936). Reading difficulties in children. *Archives of Disabled Children*, 11, 143-156.

FAGAN-DUBIN, L. (1974). Laternal dominance and development of cerebral specialisation. *Cortex*, 10, 69-74.

GORDON, H. (1921). Left-Handedness and mirror writing, especially among defective children. *Brain*, 43, 313-368.

GRANJON-GALIFRET, N., and DE AJURIAGUERRA, J. (1951). Troubles de l'apprentissage de la lecture et dominance laterale. *Encephale*, 40, 385-398.

HARDYCK, C. (1977). Handedness and part-whole relationships: a failure to replicate. *Cortex*, 13, 177-183.

HARDYCK, C., GOLDMAN, R. D., and PETRINOVICH, L. (1975). Handedness and sex, race and age. *Human Biology*, 47, 369-375.

HARDYCK, C., PETRINOVICH, L., and GOLDMAN, R. D. (1976). Left-handedness and cognitive deficit. *Cortex*, 12, 266-279.

INGRAM, T. T. S. (1959). Specific developmental disorders of speech in childhood. *Brain*, 82, 450-467.

KUTAS, M., McCARTHY, G., and DONCHIN, E. (1975). Differences between sinistrals' and destrals' ability to infer a whole from its parts: a failure to replicate *Neuropsychologia*, 13, 455-464.

LEVY, J. (1969). Possible basis for the evolution of lateral specialisation of the human brain. *Nature*, 224, 614-615.

NEBES, R. D. (1971). Handedness and the perception of part-whole relationships. *Cortex*, 7, 350-356.

NEBES, R. D., and BRIGGS, G. G. (1974). Handedness and the retention of visual material. *Cortex*, 10, 209-214.

ORTON, S. T. (1937). *Reading, Writing and Speech Problems in Children*. London: Chapman and Hall.

SATZ, P., and SPARROW, S. (1970). Specific developmental dyslexia: a theoretical formulation. In BAKKER, D. J., and SATZ, P. (Eds.), *Specific Reading Disability: Advances in Theory and Method*. Rotterdam: Rotterdam University Press.

WINER, B. J. (1971). *Statistical Principles in Experimental Design* (2nd edit.), New York: McGraw-Hill.

WUSSLER, M., and BARCLAY, A. (1970). Cerebral dominance, psycholinguistic skills and reading disability. *Percept. mot. Skills*, 31, 419-425.

(Manuscript received December 22, 1976)

Br. J. educ. Psychol., 47, 312-321, 1977

TEACHING PERSPECTIVE ABILITY TO FIVE-YEAR-OLDS

By M. V. COX
(Department of Psychology, University of York)

SUMMARY. Five-year-olds were individually trained in a structured programme of perspective problems. Compared with a control group (N=6), the children in the experimental group (N= 12) transferred their learning to both 'near' transfer tasks involving perspective-taking and 'far' transfer tasks (other concrete operational tasks). Children trained in a matrices task (N=6) transferred their learning to other 'near' transfer tasks (related matrices problems), but failed to achieve higher scores than controls in the perspectives tasks. Therefore, perspective ability requires not only an operational level of thought, but also the specific skill of adopting another's viewpoint.

INTRODUCTION

Perspective ability is the ability to represent how objects look relative to each other from another person's viewpoint. In Piaget and Inhelder's (1956) well-known 'three-mountains' task the child is seated before a model of the three mountains and asked to represent what a doll can 'see' by choosing the correct view from a selection of pictures.

Piaget and Inhelder concluded that the pre-operational child (up to about 7 years) cannot do this. He is egocentric; he cannot imagine the view from another person's position, and generally he chooses his own view saying that it represents the doll's view. At the concrete operational stage (from about 7 years) the child gradually learns to represent other views: first, he can represent the objects in the before-behind dimension correctly and later the objects in the left-right dimension. Laurendeau and Pinard's (1970) replication of Piaget and Inhelder's study generally supported their findings.

Although methodological problems have been studied extensively (eg, Garner and Plant, 1972; Cox, 1975) and the stages in the acquisition of perspective-taking skills have been further elucidated (eg, Flavell, 1974), few investigations have dealt with the *causes* of transition from one stage to another. Perhaps this is because of the difficulty of devising experiments which can monitor naturally occurring developmental changes. One approach to gaining greater understanding of developmental processes has been through attempts to accelerate normal development by training. However, as Inhelder *et al* (1974, p243) point out, 'Only if training procedures are effective can one hope to observe acquisition processes at work . . . ' Piaget himself (1964) has suggested two criteria by which the effectiveness of training may be judged: the persistence of the acquisition over time, and transfer of learning to other problems.

To date, the attempts to train perspective ability (Arond, 1966; Eliot, 1966; Miller *et al,* 1969; Miller and Miller, 1970; Iwata, 1974; Strauss and Cohen, 1974) have not been effective. They have either failed to provoke significant improvement or they have not adequately satisfied Piaget's criteria of effectiveness.

With regard to the first of these criteria, Miller and Miller (1970) tested for retention seven months after training, and Eliot (1966) after two weeks; Arond's (1966) retention test, however, was given after only two or three days. Miller *et al* (1969), Strauss and Cohen (1974), and Iwata (1974) did not report any retention tests. With regard to the second criterion of effectiveness, whereas Arond (1966) found some transfer effects to similar tasks, Strauss and Cohen (1974) did not. Eliot (1966), Miller *et al* (1969), Miller and Miller (1970), and Iwata (1974) did not employ transfer tests.

Thus, although most of these studies conclude that perspective ability was successfully induced, this is questionable. Furthermore, although control groups were included in most of the studies, only in Arond's second experiment did this group receive the same amount of experimenter time as the experimental groups. Since failure to give adequate attention to the control group may in itself account for differences in performance between experimental and control children after the training period, it is a serious design deficiency.

The training reported here was designed to develop perspective ability in pre-operational 5-year-old children. In order to assess the persistence of this ability tests were conducted four days, seven weeks, and seven months after training had ceased. Also, two types of transfer tasks were given: 'near' transfer tasks, which were variations on the perspectives task itself, and 'far' transfer tasks which included matrices tasks, a cross-sections task, and a water-level task. The control children received the same amount of experimenter time as the perspectives-trained children.

The ability to represent the relationships among a group of objects as seen from another observer's position involves two different but interrelated skills: multiple classification and perspective-taking. In order to represent the relationships among a group of objects such as the three mountains, a child must simultaneously combine two spatial relationships in order to locate each object correctly, ie, each object can be categorised as, say, 'foreground-right' or 'background-left'. This ability to combine or multiply relationships is characteristic of thinking at the concrete operational stage. In addition to the ability to classify objects in this way, perspective ability also requires the capacity to relate this to different points of view; in other words, the child must also possess perspective-taking skills.

To assess the extent to which training in multiple classification alone, *without* perspective-taking, might lead to an improvement in perspective ability, an additional group of children was trained in a matrices task (successful performance of which depends on the ability to combine two class attributes such as colour and shape). It was predicted that these children would be able to transfer their learning to other matrices problems but not to perspectives problems. On the other hand, it was predicted that the perspectives-trained children would transfer their learning to both perspectives and matrices problems.

METHOD

Experimental groups

The perspectives task and the matrices task were given to 63 children aged between 5;4 and 6;1 years, 28 of them boys and 35 of them girls. Those children who scored between 0 and 2 on both tests were retained for further testing. They were given the English Picture Vocabulary Test (Brimer and Dunn, 1962) and Raven's Coloured Progressive Matrices (Raven, 1965), to ensure that the groups would be matched in general ability, and the cross-sections and water-level tasks.

On the basis of the scores obtained with these tests, three groups of pre-operational children (N=12, N=6, N=6), closely matched in age, test scores and sex ratio, were selected. Details of the groups are shown in Table 1.

TABLE 1
DETAILS OF THE GROUPS

Perspectives group	Matrices group	Control group
N=12 6 boys 6 girls	N=6 3 boys 3 girls	N=6 3 boys 3 girls
Mean age 5;9 Range 5;6—6;0 SD 2 months	Mean age 5;10 Range 5;6—6;0 SD 2 months	Mean age 5;10 Range 5;6—6;0 SD 2 months
P score 0, N=8 1, N=2 2, N=2	P score 0, N=3 1, N=2 2, N=1	P score 0, N=3 1, N=1 2, N=2
M score 0, N=5 1, N=4 2, N=3	M score 0, N=3 1, N=2 2, N=1	M score 0, N=3 1, N=1 2, N=2
EPVT mean 112.58 Range 100-123 SD 7.60	EPVT mean 112.17 Range 97-125 SD 11.03	EPVT mean 110.83 Range 98-122 SD 9.35
RCPM mean 15.17 Range 13-18 SD 1.75	RCPM mean 15.17 Range 14-17 SD 0.98	RCPM mean 16.33 Range 13-19 SD 2.07
CS median 4.00 Range 2-8	CS median 5.50 Range 1-7	CS median 5.50 Range 1-7
WL median 5.00 Range 2-8	WL median 5.00 Range 2-7	WL median 5.00 Range 3-7

P — Perspectives task
M — Matrices task
EPVT — English Picture Vocabulary Test
RCPM — Raven's Coloured Progressive Matrices
CS — Cross-section task
WL — Water-level task

Materials and procedure

Children were tested individually by the same experimenter who also acted as the other observer in the perspectives tasks. Each task was administered in a separate session; the maximum time for each session was 25 minutes. The overall plan of the experiment is shown in Table 2.

TABLE 2
OVERALL PLAN OF THE STUDY.

PRE-TEST (1) Perspectives task (two pictures) (2) Matrices task (3) Cross-sections task (4) Water-level task (5) Raven's Coloured Progressive Matrices (6) English Picture Vocabulary Test	POST-TEST (4 days after training) (1) Perspectives task (five pictures) (2) Perspectives transfer task (different object attangement, five pictures) (3) Perspective transfer task (different positions occupied by the experimenter: 90° and 270°) (4) Matrices task (5) Matrices transfer task (different position of the empty cell) (6) Matrices transfer task (intersections) (7) Cross-sections task (8) Water-level task
TRAINING 3 conditions: perspectives-trained group, matrices-trained group, control group 20 sessions, 15 minutes each, 2 school-days apart	1st RETENTION TEST (7 weeks after training) Repeat of the post-test; administered by an independent experimenter 2nd RETENTION TEST (7 months after training) Repeat of the post-test

None of the materials in the test trials was used in the training procedures. Practice trials were given in each task.

Pre-test

(1) *Perspectives task (two pictures).* The child was seated at a round table on which was placed a black circular baseboard. Chairs were placed around the table at 90°, 180° and 270° to the child. The experimenter sat opposite the child (180° position).

In the practice trial, the experimenter placed three cardboard mountains of equal height but different colour on the baseboard. These were arranged in a triangle as in Piaget and Inhelder's experiment.

Two pictures representing the child's view and the experimenter's view of the mountains were placed in front of the child. Each picture consisted of cut-out shapes representing the mountains and the baseboard pasted on to a white background. The colours in the pictures corresponded exactly with those of the mountains. The experimenter explained that the pictures represented 'how the mountains look from different places'.

The child was asked to 'choose the one that matches how you see the mountains from your place'. He was then asked to choose the picture which corresponded to the experimenter's view, 'Which one will match how I see them from my place?' Whether or not the child made the correct choice, he checked his response by taking the picture to the experimenter's position: when in this position he was asked to decide for himself whether or not his choice was correct. If he was unable to do this, the experimenter helped by encouraging him to judge if each object was in the same relative position in the picture as in the object array. If his choice was incorrect he was asked to explain why it was wrong and he tried again until he succeeded. The experimenter pointed out the spatial relationships among the mountains as seen from her position.

A similar procedure was followed for the test trials. There were 10 object arrays: balls, bells, skittles, candles, funnels, light-bulbs, and four types of bottles. The order of these arrays was randomised for each child. Pairs of pictures were presented with each array. The child was not allowed to move from his position to check his response during the test trials.

(2) *Matrices task.* The child was required to complete a 2×2 matrix. The matrices were divided into four cells. The lower-right cell was empty. Any two adjacent cells were related in terms of either colour or shape. These attributes were randomly allocated to rows or columns. Five response cards were presented with each matrix; they included three cards identical with the three filled cells, a card with only one attribute correct for the empty cell, and the correct card which had both attributes correct.

In the practice trial the child was asked to explain his response. If he made an incorrect choice, or did not give a full explanation, the experimenter asked him first to state the connection between the items in the upper row and then the connection which should exist between items in the lower row. If, for instance, the relevant attribute for the rows was shape, the child was asked to select all those cards showing the correct shape; he was asked to state the connection between the two items in the left-hand column and then the connection between those in the right-hand column. The required colour would limit the choice to one card. The child completed the matrix with this card, and gave an explanation in terms of the two relevant attributes. No explanations were required from the child nor given by the experimenter in the ten test trials.

(3) *Cross-sections task.* A white cylindrical solid and an array of five cards (black shapes pasted on to white card) were presented in the practice trial. The child was required to select the cards which corresponded to the vertical and horizontal cross-sections of the cylinder. Twenty-four children aged between 5;0 and 6;6 had previously been given this task, but had been asked to draw the shape of the cross-sections. As well as the correct responses, the cards in the final task included the most frequent errors made by these children.

In the practice trial, the experimenter held a knife over the upright cylinder and asked, 'If I cut right through this, what shape would the inside be?' After the child had selected a card the experimenter showed him an identical cylinder which had already been cut. When this cylinder had been 'cut' through and the child shown the black shape of the inside, he corrected his response if necessary. The cards were rearranged and a similar procedure was followed for the horizontal cross-section.

The five objects used in the test trials were a blue four-sided pyramid, a brown half-cylinder, a grey prism, a green cone, and a red three-sided pyramid. An array of five cards was presented with each object. The order of the vertical and horizontal questions was randomised. The child received no feedback in the test trials.

(4) *Water-level task.* The child was shown a glass jar half-filled with coloured water. This was placed on a box so that the water surface was at the child's eye-level. A black paper sleeve was put over the jar. Five pictures were spread out on the table; the child was asked, 'How does the water in the jar look?' and required to point to the correct picture. The task was previously given to 24 children aged 5;0 to 6;6 who were asked to draw the water on prepared outline drawings of the jar. The pictures in the final task included the most frequent errors made by these children.

After the child had selected a picture the sleeve was removed and he corrected his response if necessary. The sleeve was then replaced, and the experimenter tilted the jar 30° to the right. Five more pictures were presented, and the child was again asked to select the correct picture. (The orientation of the jar in the pictures matched the angle of the actual tilted jar.) The child was provided with feedback and the opportunity to correct his response.

The experimenter tilted the jar a further 30° for each test trial. Five pictures were presented each time. The sleeve remained over the jar, and the child received no feedback about his responses.

(5) *Raven's Coloured Progressive Matrices.*

(6) *English Picture Vocabulary Test.*

Perspective training

There were 20 training sessions. Each child was trained individually in small, structured steps, starting very simply and gradually getting more complex; the order of introducing new elements into the training was based on the results of a previous study (Cox, 1978). The experimenter sat opposite the child at a round table. Throughout the training the child received verbal feedback from the experimenter about his performance; he was also allowed to change his position to check his response.

The child first learned to represent different views of one familiar object as seen from his own and then the experimenter's position. Next, he learned to represent two objects in a before-behind relationship; then objects in a left-right relationship. Finally, he learned to integrate these separate elements so that he could represent another's view of a three-object array in which both before-behind and left-right relationships are involved. By the end of the training the children were choosing from a selection of five pictures which included confusions of the before-behind and left-right relationships. These confusions were included so that the child would be forced to consider all the relationships among the objects in order to choose the correct picture consistently. When he was asked to choose a picture, this was handed to the experimenter who then evaluated the response; moreover, to check for himself, the child moved to the experimenter's position. Another picture was chosen if necessary.

Matrices training

Each child received 20 individual training sessions. The experimenter sat beside the child, who was encouraged to give explanations for his responses. The child learned how to sort a pile of plastic pieces according to two criteria: shape (rings and squares) or colour (red and yellow). Having made the first division (say, according to shape), he was then asked to sub-divide each pile (by colour). A 2×2 matrix was formed from the four piles and the child learned the relationship between the adjacent cells in each row and column. Gradually the items in each cell were reduced to one. The child learned to alter an incorrect matrix by choosing the correct item and to complete a matrix correctly by filling the empty cell. Subsequently, pictures were substituted for the pieces of plastic. At first only two response pictures were presented, but later more were added until the child had to choose from five cards.

Training for the control group

Children in the control group played '10 questions' with the experimenter using a sheet of pictures which could be divided and sub-divided into smaller categories in such a way that it was always possible to get the correct answer in the 10 questions, provided the questions were asked in an appropriate order. The child and the experimenter took turns to think of a picture and to ask the questions. In the course of the 20 sessions, the number of pictures was increased and the categories became more abstract.

(Further details of the materials and training procedures are available from the author.)

Post-test

(1) *Perspectives task (five pictures).* The same as the pre-test but five pictures were presented: the experimenter's view, the experimenter's view with only the nearest object correctly positioned, the experimenter's view with only the front and back objects correctly positioned, the experimenter's view with only the left and right objects correctly positioned, and the child's own view.

(2) *Perspectives transfer task (different object arrangement, five pictures).* The same as the pre-test but the object array was reversed and five pictures (as above) were presented with each array.

(3) *Perspectives transfer task (different positions occupied by the experimenter: 90° and 270°).* The same as the pre-test but the experimenter occupied the 90° position for five of the test trials and the 270° for the other five; the order of these was randomised. Four pictures were presented in each trial; these were the same as those in the first perspectives task of the post-test but excluded the experimenter's view with only the nearest object correctly positioned.

(4) *Matrices task.* The same as the pre-test.

(5) *Matrices transfer task (different position of empty cell).* The same as the

matrices task, but the matrix was rotated so that the empty cell was in the upper-left position.

(6) *Matrices transfer task (intersections).* The same procedure as the matrices task, but instead of 2×2 matrices, intersections were used. Three cells on the vertical axis intersected with three cells on the horizontal axis. The common cell was in the lower-right corner. Colour and shape were the relevant attributes, and were randomly allocated to each axis. Five response cards were presented with each interesection, and the child was asked to choose the correct card for the empty cell.

(7) *Cross-sections task.* The same as the pre-test.

(8) *Water-level task.* The same as the pre-test.

1st Retention test. The same as the post-test.

2nd Retention test. The same as the post-test.

RESULTS

Because perspective ability is the ability to represent views different from one's own, performance was assessed by the number of correct responses given by the child when asked to represent the *experimenter's* view. Selection of a picture in which the left and right objects were reversed would indicate the ability to represent the before-behind dimension of another's view. However, in order to ensure that the children in the experimental group could integrate before-behind *and* left-right relationships, and to provide a more stringent criterion for the existence of an operational level of thought, only correct responses were scored. Similarly only correct responses were scored in the other tasks. The maximum possible score for each task was 10. Practice trials were not scored. Table 3 shows the median correct response scores for each of the tasks.

TABLE 3
MEDIAN CORRECT RESPONSE SCORE FOR EACH TASK.

Task	Group	Pre-test	Post-test	1st Retention test	2nd Retention test
Perspectives task	P	0	10.00	9.00	7.00
	M	1.00	2.00	0.50	1.50
	C	1.00	0	0	0
Perspectives transfer task (1)	P	—	8.50	9.00	7.00
	M	—	0	1.00	1.00
	C	—	0	0	0
Perspectives transfer task (2)	P	—	7.00	8.00	8.50
	M	—	1.00	1.00	2.00
	C	—	0	1.00	1.50
Matrices task	P	1.00	9.00	9.00	10.00
	M	0.50	10.00	10.00	10.00
	C	0.50	4.00	4.50	8.00
Matrices transfer task (1)	P	—	9.00	9.50	10.00
	M	—	9.00	10.00	10.00
	C	—	6.50	2.00	8.00
Matrices transfer task (2)	P	—	9.00	9.00	9.50
	M	—	10.00	10.00	8.50
	C	—	2.00	5.00	3.50
Cross-sections task	P	4.00	5.50	6.00	5.50
	M	5.50	6.00	6.00	5.50
	C	5.50	6.00	5.50	6.00
Water-level task	P	5.00	7.00	7.00	6.50
	M	5.00	5.00	5.50	7.50
	C	5.00	3.50	5.00	5.50

P—perspectives-trained group. M—matrices-trained group. C—control group.

TABLE 4

DIFFERENCES AMONG THE GROUPS IN EACH TASK.

	PRE-TEST			POST-TEST			1st RETENTION TEST			2nd RETENTION TEST		
	P v M	P v C	M v C	P v M	P v C	M v C	P v M	P v C	M v C	P v M	P v C	M v C
Perspectives task	H=0.72, 2df, NS			H=16.31, 2df, P<.001			H=17.44, 2df, P<.001			H=10.54, 2df, P<.01		
				<.01	<.001	NS	<.001	<.001	NS	<.05	<.01	NS
Perspectives transfer task (1)				H=15.70, 2df, P<.001			H=10.52, 2df, P<.01			H=7.18, 2df, P<.05		
				<.01	<.001	NS	<.05	<.01	NS	NS	<.05	NS
Perspectives transfer task (2)				H= 8.20, 2df, P<.001			H=12.17, 2df, P<.01			H=8.50, 2df, P<.05		
				<.01	<.001	NS	<.01	<.01	NS	<.05	<.05	NS
Matrices task	H=0.16, 2df, NS			H=6.48, 2df, P<.05			H=9.79, 2df, P<.01			H=7.43, 2df, P<.05		
				NS	NS	<.05	NS	<.01	<.01	NS	<.05	<.05
Matrices transfer task (1)				H=1.86, 2df, NS			H=7.97, 2df, P<.05			H=1.42, 2df, NS		
							NS	<.05	<.05			
Matrices transfer task (2)				H=6.79, 2df, P<.05			H=10.35, 2df, P<.01			H=10.08, 2df, P<.01		
				NS	<.05	<.05	NS	<.01	<.01	NS	<.01	<.05
Cross-section task	H=0.99, 2df, NS			H=0.20, 2df, NS			H=0.81, 2df, NS			H=0.30, 2df, NS		
Water-level task	H=2.39, 2df, NS			H=4.91, 2df, NS			H=4.40, 2df, NS			H=0.58, 2df, NS		

P—perspectives group. M—matrices group. C—controls.

Because the data did not satisfy the criteria of normality of distribution and homogeneity of variance, non-parametric statistical procedures were used. The tests were the Kruskal-Wallis one-way analysis of variance by ranks, yielding an H value (Siegel, 1956, pp184-193), and Dunn's test of multiple comparisons using rank sums (Dunn, 1964). Table 4 shows the results of the statistical analysis.

The scores in the perspectives tasks of the children in the perspectives-trained group were significantly higher than those of controls. This difference was maintained after seven months. At the post-test (four days after training) the children in the perspectives-trained group achieved higher scores than the controls in only one of the three matrices tasks. However, significant differences between these two groups were observed in all three tasks after seven weeks and in two of the tasks after seven months. No significant differences were observed between the perspectives-trained group and the control group in the cross-sections and the water-level tasks.

After training, the children in the matrices-trained group attained significantly higher scores than controls in the matrices task and in one of the matrices transfer tasks. After seven weeks, their scores were significantly higher in all three of the tasks, and after seven months significantly higher in two of them. The matrices-trained group and the controls were not differentiated by their performance in any of the perspectives tasks, nor in the cross-sections and water-level tasks.

Whereas the scores of children in the perspectives-trained group were significantly higher than those in the matrices-trained group in the perspectives tasks, there were no significant differences between these groups in any of the other tasks.

DISCUSSION

The perspectives-trained children achieved higher scores than control children in both the perspectives task and the two perspectives transfer tasks; moreover, this difference was maintained seven months after training had ended. Thus, the perspectives-trained children transferred their learning to 'near' transfer tasks and this effect was lasting. Also, some transfer of learning to 'far' transfer tasks was demonstrated: the perspectives-trained children achieved significantly higher scores than controls in some matrices tasks and maintained their superiority over time; however, no significant differences were observed between the groups in the cross-sections and the water-level tasks.

Since the criteria of transfer of learning and persistence over time were generally met, it is concluded that the training had been effective.

Although the matrices-trained children achieved higher scores than controls in the matrices tasks and, to a lesser extent, in the matrices transfer tasks, there were no significant differences between these two groups in the perspectives tasks. Training in multiple classification alone does not therefore appear to be a sufficient preparation for successful performance in a perspectives task; for this, the ability to adopt other viewpoints seems to be needed as well as the ability to combine spatial relationships.

Since neither of the trained groups transferred their learning to the cross-sections and water-level tasks, it seems likely that spatial ability depends not only upon the basic requirement of an operational level of thought, but also on the possession of skills specific to each type of spatial task.

Piaget (in Weinheimer, 1972) has questioned the desirability of attempting to accelerate development. In the present study, however, the aim was to devise a training programme which could be used to study the causes of transition from one stage of development to the next. In other words, the training was used as a method of research rather than as a means of accelerating development for its own sake.

One of the aims of research is to discover the causes of normal development; consequently, a major consideration in acceleration studies is the extent to which the process of acquiring a concept through training reflects the *normal* development of that concept. Since the sequence of steps in the present study was based on the order of acquisition of perspective ability ascertained previously (Cox, 1978), the programme used approximates to normal development. The effectiveness of such a programme, as demonstrated here, emphasises the importance of teaching new concepts in a structured and problem-orientated way.

REFERENCES

AROND, N. (1966). Acceleration in the acquisition of perspective. Unpublished Doctoral Dissertation, University of Minnesota.

BRIMER, M. A., and DUNN, L. M. (1962). *Manual for the English Picture Vocabulary Tests.* Bristol: Educational Evaluation Enterprises.

COX, M. V. (1975). The other observer in a perspectives task. *Br. J. educ. Psychol.,* 45, 83-85.

COX. M. V. (1978). Order of acquisition of perspective-taking skills. *Dev. Psychol.* (in press).

DUNN, O. J. (1964). Multiple comparisons using rank sums. *Technometrics,* 6, 241-252.

ELIOT, J. (1966). The effects of age and training upon children's conceptualization of space. Ed. D. Thesis, Stanford University.

FLAVELL, J. H. (1974). The development of inferences about others. In MISCHEL, T. (Ed.), *Understanding Other Persons.* Oxford: Blackwell.

GARNER, J., and PLANT, E. L. (1972). On the measurement of egocentrism: a replication and extension of Aebli's findings. *Br. J. educ. Psychol.,* 42, 79-83.

INHELDER, B., SINCLAIR, H., and BOVET, M. (1974). *Learning and the Development of Cognition.* London: Routledge and Kegan Paul.

IWATA, J. (1974). The effect of sensory-motor cues on transformation of spatial representation in children. *Jap. J. educ. Psychol.,* 22, 21-31.

LAURENDEAU, M., and PINARD, A. (1970). *The Development of the Concept of Space in the Child.* New York: International Universities Press.

MILLER, J. W., BOISMIER, J. D., and HOOKS, J. (1969). Training in spatial conceptualization: teacher directed activity, automated and combination programs. *J. exp. Educ.,* 38, 87-92.

MILLER, J. W., and MILLER, H. G. (1970). A successful attempt to train children in coordination of projective space. Paper read at American Educational Research Association Conference, Minneapolis, March, 1970.

PIAGET, J. (1964). Cognitive development in children. In RIPPLE, R. E., and ROCKCASTLE, V. N. (Eds.), *Piaget Rediscovered.* Ithaca, New York: Cornell University Press.

PIAGET, J., and INHELDER, B. (1956). *The Child's Conception of Space.* London: Routledge and Kegan Paul.

RAVEN, J. C. (1965). *Guide to Using the Coloured Progressive Matrices.* London: H. K. Lewis.

SIEGEL, S. (1956). *Nonparametric Statistics.* New York: McGraw Hill.

STRAUSS, S., and COHEN, J. (1974). A study of the relations between mental operations underlying rotation and perspective tasks. Unpublished manuscript, Tel-Aviv University.

WEINHEIMER, S. (1972). How to teach your child to think: a conversation with Jean Piaget, the 'giant in the nursery', *Redbook Magazine,* March, 96-97, 118, 120.

(Manuscript received February 21, 1977)

Br. J. educ. Psychol, 47, 322-326, 1977

ON PRESENTING PICTURES AND SENTENCES: THE EFFECT OF PRESENTATION ORDER ON SENTENCE COMPREHENSION IN NORMAL AND MENTALLY HANDICAPPED CHILDREN*

By K. WHELDALL* and P. MITTLER
(Hester Adrian Research Centre, University of Manchester)

SUMMARY. Thirty-nine normal pre-school and 48 older ESN(S) children of similar vocabulary age were tested on a sentence comprehension test under three different presentation conditions. The test requires the child to demonstrate his comprehension of spoken sentences by pointing appropriately to one of three or four alternative visual referents. In the first condition the pictures were presented prior to the stimulus sentence, in the second the pictures and sentence were presented simultaneously, and in the third the sentence was presented prior to the pictures. Presentation conditions did not affect performance significantly for normal children, but that of ESN(S) children was adversely affected by the condition in which sentences were presented prior to pictures.

INTRODUCTION

Many tests which require the testee to perform manipulations upon materials assume that failure to do so indicates *inability* to do so. Incorrect responses, however, may arise as a result of failure to understand the instructions. Successful comprehension of instructions, moreover, may be heavily dependent upon test conditions, including the different orders of presentation of materials and instructions. In the Stanford Binet (Terman and Merrill, 1960), for example, three different presentation orders are included. 'Comparison of Sticks' requires the presentation of visual followed by verbal commands. This may be characterised as Visual/Auditory (V/A) presentation. Simultaneous (S) presentation of auditory and visual information is required in the subtest 'Identifying Parts of the Body' whereas 'Memory for Designs' is dependent upon presentation conditions in which verbal instructions are presented prior to the visual material, Auditory/Visual presentation (A/V).

It may be argued that it matters little, from a methodological point of view, which order of presentation is used, providing the order is specified and kept constant in all test administrations. It is important, however, in terms of psychological theory underlying the nature of comprehension and the models we construct of its operation; it is especially important to determine optimal conditions for comprehension when it is not comprehension in which we are primarily interested but another cognitive operation, the success of which is dependent upon successful comprehension of instructions. In the case of mentally handicapped children, whose receptive skills are more limited, it is particularly important to optimise the conditions for successful comprehension.

The Sentence Comprehension Test (Hobsbaum and Mittler, 1971) used in this study adopts a simplified model of comprehension whereby the experience/environment or non-verbal referent of the verbal message is restricted to a choice of three or four visual referents providing alternative grammatical interpretations of the verbal input, of which only one fits the verbal message perfectly. Thus the child has to scan up to four pictures which differ from each other in only one essential aspect. At the same time the child has to analyse the verbal input and eventually match it to the appropriate pictorial referent. SCT thus provides a suitable means for examining the influence of varied presentation conditions on sentence comprehension.

In comparing normal and mentally handicapped children at a similar level of receptive language development, we hypothesised:

(a) that varying presentation conditions would be more likely to influence adversely the performance of the subnormal children;

(b) that they would experience particular difficulty in the simultaneous condition, since this requires the child to carry out a simultaneous auditory and visual analysis and to relate what he has heard to what he has seen.

---

*Now at Department of Educational Psychology, University of Birmingham

METHOD

Sample

The subjects consisted of 39 normal nursery school children (mean CA 4 years 5 months) and 48 severely educationally subnormal children attending special schools (mean CA 12 years 6 months). The mean EPVT vocabulary age of the nursery school children was 4 years 2 months (SD 6.90 months) which was not significantly different from that of the ESN(S) children (3 years 11 months, SD 7.45 months).

Design

Both normal and ESN(S) children were selected on the basis of their vocabulary ages being within the range of the pre-school version of the EPVT (ie, 3 years to 4 years 11 months) and were randomly assigned to one of three experimental groups, apart from ensuring an equal sex ratio. The three normal and three ESN(S) groups formed in this way were thus at a similar level of receptive vocabulary development insofar as their mean vocabulary ages were very similar. EPVT vocabulary age has previously been shown to correlate highly with Sentence Comprehension Test score (see Mittler *et al*, 1974). Each group was tested on the Sentence Comprehension Test in one of three differing procedural forms (detailed later).

Materials

The Sentence Comprehension Test consists of a booklet of black and white picture arrays and aims to provide data on the child's ability to demonstrate comprehension of sentences of gradually increasing length and complexity of structure. Research on the test so far is reported fully in Mittler *et al* (1974) and summarised in Wheldall (1976). In brief, the child's task is to demonstrate his comprehension of a spoken sentence by correctly identifying which of three or four simultaneously presented pictures corresponds to a given spoken sentence. The pictorial referents are systematically varied to illustrate alternative grammatical interpretations.

The test consists of 15 subtests, corresponding to sentence types of varying complexity and grammatical structures including comparatives and superlatives, past and future tenses, passives, negative, plurals, prepositions and embedded clauses. Each subtest comprises four different item sentences, identical in structure, so that a more reliable assessment is possible. If a child scores three or four items correct out of four, he is said to have passed the subtest, since the probability of doing so by chance is only 5 per cent. For reasons of testing time and economy, five subtests were employed, selected so as to sample the range of difficulty of the test, as follows:

I Simple intransitive
  eg, The cat is sleeping
II Simple transitive
  eg, The girl is cutting the cake
III Simple intransitive plus adjective
  eg, The old man is reading
IV Transitive including adjective, subject, verb and object
  eg, The fat lady is pushing the pram
V Passive
  eg, The car is being pushed by the bus

Procedure

Rapport was established with each child prior to testing by means of six simple practice items (eg, Show me 'the big chair'). This also ensured that the child knew what was required of him in the test situation. Each child was then given the five subtests in the same order as previously listed and in the procedural form appropriate to the experimental condition in which he was a subject. The procedural forms were characterised as following:

Subtest 1, 'The cat is sleeping'.

V/A Condition E: "Look"—presents pictorial array—(3 second delay)—"The cat is sleeping"—"Show me". S responds.

S Condition E: "Now"—presents pictorial array and says: "The cat is sleeping"—(3 second delay)—"Show me". S responds.

A/V Condition E: "Listen, the cat is sleeping"'—(3 second delay)—presents pictorial array and says "Show me". S responds.

Note that "Look", "Now" and "Listen" were used to direct the child's attention to the start of each item. The three procedural forms serve as exaggerated models of the three testing situations previously described, in order to demonstrate the effect as clearly as possible.

RESULTS

An analysis of the data in terms of number of subtests passed, as previously defined, showed clearly the effects of presentation conditions on test performance. Table 1 shows the mean number of subtests passed for each group.

TABLE 1

MEAN NUMBER OF SUBTESTS PASSED IN THE THREE PRESENTATION CONDITIONS FOR NORMAL AND ESN(S) CHILDREN.

	Conditions		
	V/A	S	A/V
Normal (N=13 per group)	2.07	1.62	1.54
ESN(S) (N=16 per group)	1.50	1.44	0.69

One way, non-parametric analyses of variance (Kruskall-Wallis) were conducted separately for the normal and ESN(S) subjects. The normal groups' scores were not significantly different over the three conditions ($H=1.09$. $P>0.05$) but the differences between the ESN(S) groups were significant ($H=7.98$, $P<0.02$). Further analysis of the ESN(S) groups (Mann-Whitney 'U' Tests) revealed that scores in the V/A condition and the S condition were not significantly different but that both were significantly different from the A/V condition (V/A *vs.* A/V, $P<0.05$; S *vs.* A/V, $P<0.02$).

Comparisons between the normal and ESN(S) groups revealed that differences in the V/A and S conditions were not significant but that the normal group was significantly superior to the ESN(S) group in the A/V condition ($P<0.05$, one-tailed).

DISCUSSION

The experiment provided some evidence to support our general hypothesis that presentation conditions would affect the test performance of ESN(S) children more than normal children. Differences between presentation conditions were significant for the ESN(S) sample but not for the normal sample. However, we failed to substantiate our original specific hypothesis that the S condition would be the most difficult, due to severe strain on the subnormal child's attentional/processing abilities. The S and V/A conditions were seen to exert a similar effect whereas the A/V condition appeared to hinder the subnormal child's performance.

Two explanations may be advanced to account for these findings, not necessarily mutually exclusive; these are based on: (a) problems of memory and storage and (b) the possibility of comparison and checking of the two inputs.

(a) In the A/V and S conditions, one might argue that subjects would have to store the message (possibly in an uncoded form) whilst scanning the pictures, after which they could then attend to the message. This would be consistent with Ellis' (1963) views:

> "The retardates' inadequate behaviour was in part due to a stimulus trace attenuated in amplitude and duration . . . From this premise, it could be predicted that when the performance of retardates and normals was compared on some task which depended upon bridging a temporal gap, the retardates' behaviour would suffer in comparison. Moreover, when the magnitude of temporal separation between events increased, the retardates' performance would deteriorate even more".

The decline in mean number of subtests passed over the three conditions suggests that delay may constitute the crucial distinction between the experimental conditions. This relates to Graham's (1970) findings on sentence comprehension which utilised a procedure similar to that employed in SCT. The results of his experiments, in which testing was carried out using a form of the A/V presentation procedure, showed that STM was the most important factor influencing sentence comprehension in his sample of ESN(M) children. Alternatively, the present research may be taken as suggesting that Graham's finding regarding the importance of STM in the sentence comprehension ability of ESN(M) children might be an artefact of the presentation condition employed. It might be argued that his A/V type procedure emphasised the STM aspects of the experimental paradigm.

(b) If memory load alone were the principal factor involved, however, we would expect the results to decline steadily but significantly over the three conditions, whereas the V/A and S conditions were not shown to exert significantly different effects for the ESN(S) group. Where, according to the theory we should have got A/V<S<V/A, we got A/V<S=V/A. This suggests that another, possibly additional, influence is involved. Several experiments have apparently demonstrated the superiority of procedures in which the verbal message precedes the pictorial referent(s) (Clark and Chase 1972; Tversky 1975). Such studies, however, more commonly measured the latencies of adult subjects responding to written verbal messages and hence are of only limited relevance.

The main advantage of the A/V conditions appears to be that the auditory input may be decoded before the visual input is received. Olson (1972) reports an experiment where young normal subjects found it easier to verify a sentence picture pair when the spoken sentence was presented prior to the picture (A/V) than when the picture was presented prior to (and disappeared before) the sentence (V/A). He comments that this is to be expected, as it is easier to recode a picture to a sentence than vice versa, and states the general hypothesis that re-coding pictures or perceptions precedes developmentally the re-coding of sentences.

In Olson's experiment, however, the comparison involved is not the same as in the present experiment, since the pictures in our V/A condition do not disappear before the sentence is presented, but are available for decoding/matching prior to, during and after the auditory input in the V/A condition. This presumably accounts for the inferiority of the A/V condition in this experiment. The effect of being able to decode the verbal message prior to presentation of pictorial referents is apparently negligible compared to the strain it imposes on short-term memory. The V/A condition imposes no similar strain, as the visual information is available for scanning/decoding throughout and so the problems of memory decay (of the visual information) before the response is demanded are absent. It is to be noted that this sort of V/A condition is the most usual comprehension situation found in real life.

In both the S and V/A conditions the memory/storage problems are reduced. Moreover, the visual material is available for comparison and checking contemporaneously with the auditory input. Given that the results for the ESN(S) group were A/V<S=V/A, in line with this view, our final explanation requires consideration of both of our previously proposed explanations. Reduced memory load, plus the chance to relate what is being heard to what is being seen, appears to improve performance for ESN(S) children.

In summary, our findings suggest that whereas presentation conditions are less likely to influence the normal child's performance, we must be cautious in interpreting the results of tests on subnormal children. If a subnormal child fails a cognitive task, it may not be due solely to his inability to carry out that task: he may have been unable to decode or retain the instructions due to the nature of the presentation conditions. It may, therefore, be appropriate to advise testers to consider experimental variations on presentation conditions before reaching definitive conclusions about the ability of subnormal children to comprehend sentences or to perform tasks dependent upon sentence comprehension.

ACKNOWLEDGMENT—The research reported in this paper was supported by SSRC grant HR 840 awarded to the second author.

REFERENCES

CLARK, H. H., and CHASE, W. G. (1972). On the process of comparing sentences against pictures. *Coq. Psychol.*, 3, 472-517.

DENNY, M. R. (1964). Research in learning and performance. In STEVENS, H. A. and HEBER, R. (Eds.), *Mental Retardation.* Chicago: University of Chicago Press.

ELLIS, N. R. (1963). The stimulus trace and behavioural inadequacy. In ELLIS, N. R. (Ed.), *Handbook of Mental deficiency,* New York: McGraw Hill.

GRAHAM, N. C. (1970). The language of educationally subnormal children. Final report. School of Education, University of Birmingham.

HOBSBAUM, A., and MITTLER, P. (1971). *Sentence Comprehension Test* (Experimental Edition). Hester Adrian Research Centre, University of Manchester.

MITTLER, P., JEFFREE, D., WHELDALL, K., and BERRY, P. (1974). Assessment and remediation of language comprehension and production in severely subnormal children. Final Report to Social Science Research Council, Hester Adrian Research Centre, University of Manchester.

OLSON, D. (1972). Language use for communicating, instructing and thinking. In CARROLL, J. B. and FREEDLE, R. O. (Eds.), *Language Comprehension and the Acquisition of Knowledge.* New York: Wiley.

TERMAN, L. M., and MERRILL, M. A. (1960). *Stanford-Binet Intelligence Scale: Manual for the Third Revision.* Boston: Houghton and Mifflin.

TVERSKY, B. (1975). Pictorial encoding of sentences in sentence-picture comparisons. *Quart. J. exp. Psychol.*, 27, 405-410.

WHELDALL, K. (1976). Receptive language development in the mentally handicapped. In BERRY, P. (Ed.), *Language and Communication in the Mentally Handicapped.* London: Arnold.

(Manuscript received February 7, 1977)

Research Notes

Br. J. educ. Psychol., 47, 327-329, 1977

STRATEGIES CHILDREN USE TO CO-ORDINATE PERSPECTIVES AS A FUNCTION OF TASK DEMANDS

By CHRISTINE EISER

(Department of Psychological Medicine, Hospital for Sick Children, London)

SUMMARY.Previous research has suggested that children are aware of changes in *orientation* prior to changes in *position* of objects as a function of an observer's viewpoint. Two question-asking studies are reported in which this finding was investigated in relation to (a) changes in the type of stimuli and (b) changes in the response mode. In a rotation task, 40 Canadian elementary school children anticipated perspective by determining the orientation of objects even where the stimuli were such as to make a focus on position more appropriate. However, where the response mode involved placing objects to match an observer's view with a sample of 30 similar children, they tended to ask questions about the *position* rather than the *orientation* of objects. Results in previous co-ordination of perspective tasks are discussed in relation to these findings.

INTRODUCTION

Recently, Eiser (1976) investigated the information children felt was needed to solve co-ordination of perspectives tasks (cf. Piaget and Inhelder, 1956). Using a technique based on the '20 questions' game (Mosher and Hornsby, 1966), she showed 7-, 9-, and 11-year-olds a model of three objects, and asked them to match their own perspective to that of the experimenter, who had a duplicate array, rotated to a new position and shielded from the child. Overall, more questions were asked about the *orientations* of the individual stimuli, but with increasing age, there were more questions about their *positions*. This supports the conclusions of Coie *et al* (1973) that children are first aware of *what* is seen from different perspectives, then of *how*, and finally of *where* objects appear to different observers. Both studies used stimuli (model houses or animals) with clearly differentiated sides such that orientation information would help solve the task. The present report examines if this preference for orientation information persists for stimuli where such information is irrelevant, and for alternative response modes.

EXPERIMENT I—METHOD

Sample

Forty children were involved; 10 boys and 10 girls in a younger group (mean age, 8 years 5 months) and 9 boys and 11 girls in an older group (mean age, 11 years 4 months). All children attended a middle-class village school in southern Ontario. Testing was carried out in a single 20-minute session for each child.

Apparatus and procedure

Two sets of stimuli were used, each consisting of three objects glued to a revolving tray. The first ('animals') consisted of the same objects (mushroom, mouse and snail) described in Eiser (1976). The second consisted of three coloured 'cones'; brown, pink and blue and measuring 17, 10 and 9 cms respectively. The experimenter had a duplicate set of each.

The procedure is described in Eiser (1976). Children were to ask questions until they could rotate their stimuli to match the experimenter's. For each set of stimuli, there were four trials, and the accuracy of response and number and type of questions was recorded for each trial.

RESULTS

For each set, following the procedure described by Eiser (1976), subjects were divided by a median split for each age group separately into those who were *efficient*, (ie, asked few questions and/or made small errors) and those who were *inefficient* (asked more questions and/or made larger errors). A qualitative analysis of the questions was then made by categorising each question as *orientation* (eg, 'Can you see the snail's face?') or *position* (eg, 'Is the mouse on the right?') or *miscellaneous*. The percentage of questions asked by each subject in each category was then calculated.

For the 'animals' stimuli, more *orientation* (79.2 per cent) than *position* questions were asked ($F=10.1$, df=1.36, $P<.005$), but for the 'cones', more position (54.6 per cent) questions were asked. This percentage was significantly higher for *efficient* than inefficient subjects (73.7 per cent *vs* 35.4 per cent $F=12.1$, $P<.01$) and older subjects tended to ask more position questions than younger (64.8 per cent *vs* 44.3 per cent, $F=3.5$).

EXPERIMENT II—METHOD

Sample

Fifteen children (mean age, 6 years 7 months) and 15 older children (mean age, 8 years 7 months) from similar backgrounds as in Experiment I, were subjects.

Apparatus and procedure

Two types of stimuli were used. The first included a model toad, rabbit and snail, which all had clearly differentiated sides and lent themselves to orientation questions. The second set lacked, as far as possible, clues to orientation, but suggested a definite position. There were three cylinders each marked '1, 2 or 3'. As before, the experimenter had a duplicate set of each.

All children first attempted two practice trials with a mushroom and a shell. They then attempted two trials with each set of stimuli, order being balanced across subjects. Children were to ask questions until they could position their stimuli to correspond with the 'pattern' made by the experimenter.

RESULTS

Again, questions were coded into the *position* or *orientation* categories. For both stimuli, the percentage of position questions was high ('animals'=77.8 per cent, 'cones'=82.3 per cent). The number of subjects asking any orientation questions increased with age, including only seven children in the younger group and 12 in the older group ($\chi^2=4.61$, $P<.05$).

DISCUSSION

Children of both ages tended to ask questions about the *position* rather than the *orientation* of objects when asked to place objects so as to correspond with another's perspective. This occurred for both types of stimuli used, suggesting the critical factor in determining the child's choice of questions is the response mode, rather than intrinsic properties of the stimuli. Thus, in the first experiment, where children are asked to rotate objects to match a display they tend to ask for information about the directions objects face; in the second experiment, where they are asked to place the objects to match another's perspective, they ask almost exclusively for information about position.

It must be noted, however, that although they asked for position information, children in the second experiment tended to assume the orientation of objects. Thus, having established an object's position, they often tended to place it so that it faced toward them, without attempting to establish that this was correct. It was only the older subjects who were concerned about how objects should be orientated.

Our results suggest that children may adopt different approaches to co-ordination of perspectives tasks as a function of task demands. For example, in rotation tasks, as described by Fishbein *et al* (1972) they may focus on the orientation of specific objects. In tasks of the kind used by Flavell *et al* (1968) they may concentrate on predicting changes in the positions of objects. In discussing the discrepancies between the ages at which children are reported to be able to co-ordinate perspectives, Flavell (1971) argued that such tasks represent very different levels of the *same* perspective-taking skill. Our own data suggest that children approach apparently similar co-ordination of perspectives tasks using different cognitive strategies. The apparent *décalage* which is often noted in such research (cf. Laurendeau and Pinard, 1970) may be at least partly attributable to qualitative differences in task demands between the paradigms typically employed.

ACKNOWLEDGMENTS—This research was submitted in partial fulfillment for the degree of Ph.D. at the University of Bristol while the author was in receipt of an SSRC scholarship. The author would like to thank Dr N. H. Freeman for advice and encouragement, and Miss P. Sorene for typing the manuscript.

REFERENCES

COIE, J. D., CONSTANZO, P. R., and FARNILL, D. (1973). Specific transitions in the development of spatial perspective-taking ability. *Dev. Psychol.*, 9, 167-177.

EISER, C. (1976). Questions children ask about spatial arrays: an analysis of the processes involved in co-ordinating perspectives. *Br. J. educ. Psychol.*, 46, 203-211.

FISHBEIN, H. D., LEWIS, S., and KEIFFER, K. (1972). Children's understanding of spatial relations: co-ordination of perspectives. *Dev. Psychol.*, 7, 21-23.

FLAVELL, J. H. (1971). The development of inferences about others. Paper presented at the Interdisciplinary Conference on our Knowledge of Persons: Person Perception and Interpersonal Behaviour, State University of New York, Binghampton, New York, December 13-16.

FLAVELL, J. H., BOTKIN, P. T., FRY, C. L., WRIGHT, J. W., and JARVIS, P. E. (1968). *The Development of Role-Taking and Communication in Children.* New York: Wiley.

LAURENDEAU, M., and PINARD, A. (1970). *Development of the Concept of Space in the Child.* New York: International Universities Press.

MOSHER, F. A., and HORNSBY, J. R. (1966). On asking questions. In BRUNER, J. S. (Ed.), *Studies in Cognitive Growth.* New York: Wiley.

PIAGET, J., and INHELDER, B. (1956). *The Child's Conception of Space.* London: Routledge and Kegan Paul.

(Manuscript received May 16, 1977)

Br. J. educ. Psychol., 47, 330-334, 1977

AN AUSTRALIAN EVALUATION OF THE JUNIOR EYSENCK PERSONALITY INVENTORY

By B. C. HANSFORD and H. NEIDHART

(University of New England, Australia)

SUMMARY. An Australian sample (N=463) of 15-year-old girls was used in an evaluation of the Junior Eysenck Personality Inventory (JEPI). In particular, this study examined the terminology used in various items, response rates, factorial structure, and the relationship between inventory scales. Although the findings regarding item terminology and response rates may indicate a need to modify JEPI, the factorial structure and relationship between scales support the Eysenckian position.

INTRODUCTION

On the basis of various factor analytic studies, Eysenck and his colleagues have built up impressive support for the contention that in the study of human behaviour there exist two important and distinct dimensions of personality. These behavioural traits are described as extraversion (E)/introversion, and neuroticism (N)/stability.

With regard to the study of personality in adults, Eysenck has developed two widely used inventories: the Maudsley Personality Inventory (MPI) (Eysenck, 1959) and the Eysenck Personality Inventory (EPI) (Eysenck and Eysenck, 1964). This paper is concerned with a development from these inventories, the Junior Eysenck Personality Inventory (JEPI) (S. B. G. Eysenck, 1965). The JEPI contains 60 items, and was designed for use with children aged seven to 16. Of the 60 items in the inventory, 24 are presented as measuring E, 24 as measuring N, and 12 as constituting a lie scale (L). Considerable information relating to standardisation and reliability of JEPI are reported in the manual (S. B. G. Eysenck, 1965) and in *Personality Structure and Measurement* (Eysenck and Eysenck, 1969).

Despite considerable empirical evidence, and the strong theoretical grounds presented for the Eysenckian argument, there have been a number of studies relatively critical of the particular model of personality. In discussing JEPI, Chazan (1971, p204) makes the following statement, "estimates of a child's personality characteristics made on the basis of self-ratings must be viewed with caution". Bennett and Youngman (1973, p228) suggest "there is little evidence that enables an accurate estimate to be made of what a JEPI score means in the school setting".

In the current study, data were collected in five girls' schools located in the metropolitan area of Sydney. It was decided to restrict the sample to 15-year-old girls (N=463). This restriction made it possible to compare results with those for one of the age groups reported in the JEPI manual (S. B. G. Eysenck, 1965).

This study examines four aspects of JEPI that have been subject to criticism, namely the terminology used in certain JEPI items, item response rate ceilings, factorial structure of the inventory, and the relationship between inventory scales.

RESULTS AND DISCUSSION

Item Analysis

The percentage response to each item was calculated. Of the 24 items making up the E scale, 11 were found to have response rates in excess of 80 per cent (Table 1). By way of contrast only one item (26) in the N scale exhibited this apparent lack of ceiling.

S. B. G. Eysenck(1965) indicated that the distribution on the extraversion scale is 'roughly normal', yet the mean score usually approaches 18 in a scale consisting of 24 items. In the current study, an examination of response rates seems to point to the possibility of certain E items lacking discrimination. Bennett (1973) provides similar evidence regarding the E scale. In Bennett's study, 12 E items were reported as having response rates greater than 80 per cent. As a result of this finding Bennett advocated modification of the E scale "to overcome the lack of discrimination of many of the items" (1973, p137). Entwistle (1974) has suggested that response rates of 80 per cent are generally viewed with concern in Likert-type tests. He also points out that the 12 items identified by Bennett generally refer to social extraversion, and that three specifically relate to behaviour at parties. In the current study the 11 E items deal with

TABLE 1

PERCENTAGE RESPONSE RATES

JEPI Item No.	Percentage Response Yes	No
1	95	5
6	8	92
11	87	13
22	95	5
25	91	9
27	97	3
38	89	11
46	92	8
51	13	87
53	81	19
59	11	89

'sociability' (three with party behaviour). This raises two possibilities. Firstly, in the translation of items from adult inventories to one suitable for children, some may have lost their meaning. Secondly, in distinguishing between extraverts and introverts the JEPI may be relying on 'impulsivity' rather than 'sociability' items.

Terminology

A series of studies conducted in the United Kingdom (Entwistle and Welsh, 1969; Eysenck and Cookson, 1969, 1970; Finlayson, 1970) have indicated the JEPI to be a suitable instrument for use with children above 10 years of age. Despite the apparent simplicity and clarity of language there are a number of questions that are difficult to answer with a simple 'yes' or 'no'. Chazan (1971) has also suggested that the frequent use of the words 'often' and 'sometimes' in the formulation of inventory items leads to possible interpretation problems.

Hall (1969), Cookson (1970), and Best (1973) have reported difficulties with particular expressions used in JEPI. Items using expressions such as 'moody', 'thumping in your heart', 'pranks', 'happy-go-lucky', and 'dizzy turns' raised questions in previous research, and during administration in the current study. Best (1973) discovered respondents adopted a variety of strategies when unsure of item meaning, and in some cases provided meaningless answers.

In the current study, Item 28—"Can you usually let yourself go and enjoy yourself at a gay party?" was re-worded, deleting the word 'gay'. This deletion followed a discussion of item suitability with several 14-16-year-old girls. The current connotation attached to 'gay' was obviously the source of considerable amusement.

Factor Analysis

A varimax solution (Kaiser, 1960) was used to rotate factors to a simple solution, and by applying the scree test (Cattell, 1966) it appeared appropriate to extract factors with eigenvalues >2.0. It should be noted that this was the same level of eigenvalue used by Bennett (1973) in his factoring of JEPI, and by Eysenck and Eysenck (1969) in a refactoring of the EPI.

The four factors retained (Table 2) in the current study accounted for 24.3 per cent of variance. This compares with the Bennett (1973) study in which five factors accounting for 26.5 per cent of the variance were retained.

Table 2 reports only three of the four retained factors, as it was felt unnecessary to present a factor that so clearly represented 'neuroticism'. This particular factor contained the 24 items comprising the neuroticism scale, all with loadings in excess of .27. The items with the highest loadings were; 31—"Do you often feel fed-up?" (.58); 52—"Do you find it hard to get to sleep at night because you are worrying about things?" (.55); 23—"Do you get tired for no good reason?" (.54).

The items loading on Factor A are mainly (12 out of 13) drawn from the E scale. An examination of these extraversion items suggests what S. B. G. Eysenck (1969, p282)

TABLE 2

ORTHOGONAL FACTOR STRUCTURE OF JEPI

Item No.	JEPI Designation		Loading
		Factor A: 'Impulsivity'	
49	E	Being very lively	.55
55	E	Shy of speaking	—.53
14	E	Have to act quickly	.51
17	E	Got a party going	.51
3	E	A quick answer	.47
33	E	Most quiet	—.46
41	E	Happy-go-lucky	.45
19	E	Make new friends	.45
11	E	Rather lively	.41
53	E	Can do the things	.40
*43	E	A rough and tumble	.34
32	L	Boast a little	.34
22	E	Like telling jokes	.25
		Factor B: 'Sociability'	
38	E	Enjoy a party	—.57
1	E	Excitement going on	—.56
59	E	A lively party	.54
6	E	Rather be alone	.53
46	E	Going out a lot	—.46
27	E	Mixing with others	—.43
51	E	Rather sit and watch	.43
35	E	Do things suddenly	—.35
		Factor C: 'Lie'	
8	L	Always do as told	.60
12	L	Ever broken rules	—.58
48	L	Ever cheeky	—.56
36	L	Always quiet	.52
24	L	Always finish homework	.49
57	E	Shout back	—.45
20	L	Ever told a lie	—.43
4	L	Sometimes get cross	—.41
28	L	Say prayers	.41
40	L	Said anything nasty	—.40
*43	E	A rough and tumble	—.40
44	L	Always eat everything	.24

*Reported in more than one factor.

calls 'social shyness' (Items 55, 3, 33, and 19). Bennett (1973) described a number of these items (49, 14, 17, 3, 19, 11 and 53) as reflecting a dimension termed 'Active-Passive'. Items 43 and 22 were described by S. B. G. Eysenck (1969, p282) as belonging to a dimension called 'jocularity'. Factor A contains aspects that appear to reflect 'liveliness', 'shyness', 'activity', and 'jocularity'; it consequently seems to represent that part of extraversion known as 'impulsivity'.

Factor B consists entirely of E items. An inspection of the items loading on this dimension suggest the importance attached to mixing with other people and attending parties. This particular dimension contains six items (59, 51, 38, 27, 6 and 11) retained in the Bennett (1973) study and described as defining a dimension of 'Sociable-Unsociable'. Clearly, Factor B reflects 'sociability', and it is perhaps worth noting that almost all items loading on this factor are those reported in Table 1 as 'lacking ceiling'.

Whereas Bennett showed an interaction between lie and stability items, this was not apparent in the current factorial solution. Of the 12 L items in JEPI, 10 were retained in Factor C (nine with loadings $\geq .40$). The two L items not loading on this factor were 32 and 16. Item 32—"Do you sometimes boast a little?" had its major loading on the 'impulsivity' dimension. Item 16—"Can you always keep every secret?" was one of the few JEPI items that loaded across the factors, thus suggesting a lack of factorial clarity.

Relationship Between Scales

In Table 3 a comparison is made between the Eysenck norms for 15-year-old girls and those for the current study.

TABLE 3

RELATIONSHIP BETWEEN SCALES OF JEPI (15-YEAR-OLD GIRLS)

Scales	Eysenck		Current Study (N=463)
	Sample A (N=258)	Sample B (N=63	
E with N	—.173	.030	—.089
E with L	NA	—.367	—.309
N with L	NA	—.368	—.211

NA—Not available.

The relationship between E and N has been conceptualised as uncorrelated and independent dimensions of personality. Typically this relationship has been shown to be slightly negative, but the values are generally close to zero and thus support the notion of independent scales. In the current study, the correlation between E and N was −0.089, and therefore between the two figures reported by S. B. G. Eysenck (Table 3).

A significant negative correlation (−0.309) is indicated (Table 3) between E and L. This correlation is of quite similar value to that reported by S. B. G. Eysenck (−0.367).

The correlation between N and L is reported by S. B. G. Eysenck (Table 3) as −0.368, and in the current study the equivalent figure was −0.211.

With regard to 15-year-old girls, the correlations obtained in this study between JEPI scales are quite similar to the data reported in the manual (S. B. G. Eysenck, 1965).

CONCLUSION

The results of this study do indicate certain problems with terminology used in JEPI items, and more importantly, with an apparent lack of ceiling reflected in response rates to specific E items.

From the results obtained in the factor analysis of responses it would seem that N and L items cluster in such a way as to give considerable support to the Eysenckian position. With regard to extraversion, this study provides supporting evidence for those who have advocated the dual nature of extraversion, namely 'impulsivity' and 'sociability'.

Entwistle (1972, p138) suggests that with respect to children the correlation between extraversion and neuroticism 'is generally found to be significantly negative and not zero'. This was not the case with the present study, as the negative correlation was too small to be of any real importance. In fact the correlations between JEPI scales were quite similar to those reported by S. B. G. Eysenck.

REFERENCES

BENNETT, S. N. (1973). A re-evaluation of the Junior Eysenck Personality Inventory. *Br. J. educ. Psychol.*, 43, 131-139.

BENNETT, S. N., and YOUNGMAN, M. B. (1973). Personality and behaviour in schools. *Br. J. educ. Psychol.*, 43, 228-233.

BEST, K. (1973). Children's situation responses to the JEPI: a validation study. M.A. Dissertation, University of Lancaster.

CATTELL, R. B. (1966). The scree test for the number of factors. *Mult. Beh. Res.*, 1, 245-276.

CHAZAN, M. (1971). Junior Eysenck Personality Inventory. In *The Seventh Mental Measurement Yearbook.* New Jersey: Gryphon Press.

COOKSON, D. (1970). A study of difficulties in reading and understanding the Junior Eysenck Personality Inventory. *Br. J. educ. Psychol.*, 40, 8-14.

COOKSON, D. (1973). The Junior Eysenck Personality Inventory and teacher's ratings: an investigation of unexpected results. *Br. J. educ. Psychol.*, 43, 52-57.

ENTWISTLE, N. J. (1974). Typologies of students and schools: investigations into the relationship between personality and attainment. In *Personality and Academic Progress.* Conference Readings, Association of Educational Psychologists, Institute of Education, University of London.

ENTWISTLE, N. J. (1972). Personality and academic attainment. *Br. J. educ. Psychol.*, 42, 137-151.

ENTWISTLE, N. J., and WELSH, J. (1969). Correlates of school attainments at different ability levels. *Br. J. educ. Psychol.*, 39, 57-63.

EYSENCK, H. J. (1959). *Manual of the Maudsley Personality Inventory.* London: University of London Press.

EYSENCK, S. B. G. (1965). *Manual of the Junior Eysenck Personality Inventory.* London: University of London Press.

EYSENCK, S. B. G. (1969). Personality dimensions in children. In EYSENCK, H. J., and EYSENCK, S. B. G. *Personality Structure and Measurement.* London: Routledge and Kegan Paul.

EYSENCK, H. J., and COOKSON, D. (1969, 1970). Personality in primary school children. *Br. J. educ. Psychol.*, 39, 123-130; 40, 117-131.

EYSENCK, H. J., and EYSENCK, S. B. G. (1964). *Eysenck Personality Inventory.* London: University of London Press.

FINLAYSON, D. S. (1970). A follow-up study of school achievement in relation to personality. *Br. J. educ. Psychol.*, 40, 344-348.

HALL, E. (1969). The validity of the JEPI for ESN children. *Br. J. soc. clin. Psychol.*, 8, 81-82.

KAISER, H. F. (1960). The application of electronic computers to factor analysis. *Educ. Psychol. Meas.*, 20, 141-151.

(Manuscript received November 22, 1976)

Research Notes

Br. J. educ. Psychol., 47, 335-338, 1977

THE EFFECT OF TYPE OF ITEM ON STUDENT-TEACHER INTERACTION DURING FEEDBACK OF EXAMINATION PERFORMANCE

By N. C. BOREHAM

(Department of Adult and Higher Education, University of Manchester)

SUMMARY. Seventy-three State Registered Nurses following a course in intensive care nursing were given achievement tests, and then given knowledge of their results in tutorials. Half the students took a test of the ability to recall knowledge, while the remainder of the sample took a test of the ability to apply such knowledge. Measures of teacher-student interaction during the tutorials showed that the application test stimulated more information exchange and greater student involvement in discussion. The significance of this is related to the informative role of feedback in facilitating learning.

INTRODUCTION

Providing students with feedback concerning their performance in periodic tests has been advocated as an instructional aid on the grounds that knowledge of results facilitates learning (eg, Beard, 1967; Ellis, 1967). Although in the past psychologists explained the facilitating effect of knowledge of results on learning in terms of reinforcement (eg, Pressey, 1926, 1927, 1950; Angell, 1949), today it is widely accepted that this effect is due primarily to the provision of information (Annett, 1969). Several investigators have found that a particularly effective method of giving feedback of examination results is discussion between teacher and students (Curtis and Wood, 1929; Gilbert, 1956), and two studies provide evidence that this is due to the increased opportunities for information exchange which occur in the tutorial situation. Sassenrath and Garverick (1965) compared three methods of giving students feedback of midsemester examinations: letting them check their answers against the correct ones written on a blackboard; discussion with an instructor; and assigning them to textbook material related to questions which they had answered incorrectly. The discussion group performed better on post-tests than the others. This was explained by the fact that members of this group were able to obtain more information from the discussion than those who could only check their answers against the blackboard, and even more than those assigned to textbook passages (who concentrated only on questions which they answered incorrectly, thereby missing information about questions which they might have got right only by chance). Montor (1970) gave students feedback of multiple-choice test results using a device which gave an immediate indication of whether or not each response was correct. A control group took the same test without this continuing feedback. However, after the testing sessions, the correct answers were read to *both* experimental and control groups by an instructor who offered to answer questions. Because they had received no other feedback, the intended control group initiated more discussion with the instructor, and thereby obtained more information than the experimental group regarding how and why particular responses were correct or incorrect. The outcome was higher post-test scores in the control group, attributed by Montor to the greater information exchange occurring in their discussion.

Although there has been much research into the psychometric properties of different types of test item, there appears to have been none on the suitability of different item types for stimulating informative classroom interactions. When examinations are used to give knowledge of results, they function as feedback loops linking teacher with student, and so in addition to possessing validity and reliability, they must be of types which facilitate the exchange of information. The aim of the present investigation was to compare items at different levels in Bloom's Taxonomy (Bloom *et al*, 1956) with respect to their suitability for this purpose.

METHOD

The sample comprised 73 State Registered Nurses following a course in intensive care nursing in ten hospitals throughout England and Wales. This course includes periods of supervised practice in intensive care units, together with occasional study days in the hospital's school of nursing, during which it is common for clinical teachers to base tutorials on the administration and discussion of short quizzes. Two tests were constructed for this purpose, representing the levels in the Bloom Taxonomy designated as 'knowledge' and 'application'. Both tests were of the short-answer types, and were based on the same material (the nursing care of renal failure). The knowledge test required students to remember and cite principles and facts in this domain. The application test required students to carry out the cognitive operations cited by Bloom *et al* (1956) as the constituents of 'application'. Presented with a problem to which he possesses no ready response, the student has to scan the cues presented in the problem-statement, identify the nature of the problem presented therein, recall the relevant principles and facts, and then use this recalled information to arrive at an answer. The development and validation of these tests is described elsewhere (Boreham, 1977).

In each hospital, students were randomly divided into two tutorial groups. One group took the application test, and the other took the knowledge test. To equalise teaching styles as far as possible, both the tutorials in each hospital were run by the same clinical teacher. Carry-over effects (whereby the teacher's experience of the first tutorial might affect how she behaved in the second) were balanced by giving the application test tutorial first in five hospitals, and second in the other five, the choice of hospitals being made by random process. All ten clinical teachers adopted broadly the same teaching strategy. They administered the test at the beginning of the tutorial, and then, taking the items one at a time, asked individual students to read out their answers, the teacher commenting on the correctness of each response, explaining points which had caused difficulty, and responding to any other issues raised by students. The last mentioned activity consisted mainly of answering supplementary questions and responding to students' suggestions of alternative answers to the ones the teacher had given.

The extent to which the tests stimulated teacher-student interaction which would facilitate information exchange was assessed by the researcher, who sat in on all the tutorials and performed an interaction analysis. Two measures were adopted: (1) the length of discussion generated by each test, and (2) Flanders Interaction Analysis (Flanders, 1970).

RESULTS

Flanders Interaction Analysis subdivides both teacher talk and student talk into the categories of *initiation* ('to introduce an idea or concept for the first time'), and *response* ('to amplify or react to ideas which have already been expressed'). The only major difference in the patterns of interaction generated by the two tests was a higher ratio of student initiation to response during most of the application test tutorials (Table 1). A Wilcoxon matched-pairs signed ranks test (Siegel, 1956) indicated that the difference was significant at the one per cent level (two-tailed probability). Two hospitals reversed the overall trend. However, in one (hospital A) the teacher volunteered the information that randomisation had unluckily placed all the most self-assertive students in the knowledge test group. And in the other (hospital C), the clinical teacher could not be prevented from restricting discussion of each test to precisely 15 minutes, due to misplaced zeal in attempting to achieve a well-controlled comparison. Since he set himself to cover all the items in each test in that time, he frequently discouraged student discussion, thus removing any differences which might otherwise have manifested themselves. For these reasons, the two inconsistent results do not cast serious doubt on the superiority of the application test for stimulating student participation in discussion along the lines which Flanders (1970) describes as 'expressing own ideas; initiating a new topic; developing opinions and a line of thought . . . going beyond the existing structure'.

In all hospitals except C, the clinical teachers allowed discussion of each item to continue until those present felt that all the issues which it raised had been covered. Since it was not usually possible to discuss the whole of each test in the time available, an index of amount of discussion was obtained by dividing the duration of the discussion in minutes by the number of items covered in that time. On the whole, lengthier discussion resulted from use of the application test (Table 1). A Wilcoxon matched-pairs signed ranks test (Siegel, 1956) showed that the difference was

TABLE 1

RATIO OF STUDENT INITIATION TO STUDENT RESPONSE, AND MINUTES OF DISCUSSION PER ITEM, DURING TUTORIALS

Hospital	Ratio of Initiation to Response			Minutes' Discussion per Item		
	Application Test group	Knowledge Test group	Difference*	Application Test group	Knowledge Test group	Difference†
Application Test administered first						
A	1.08:1	1.14:1	−0.06	3.8	2.5	+1.3
E	0.32:1	0.17:1	+0.15	2.5	1.8	+0.7
F	0.30:1	0:1	+0.30	2.4	1.0	+1.4
G	0.22:1	0.05:1	+0.17	2.3	2.0	+0.3
J	0.31:1	0.12:1	+0.19	4.9	3.6	+1.3
Knowledge Test administered first						
B	0.29:1	0.21:1	+0.08	2.3	1.7	+0.6
C	0:1	0.03:1	−0.03	1.3	1.3	0
D	0.93:1	0.41:1	+0.52	2.1	2.3	−0.2
H	0.22:1	0.05:1	+0.17	2.3	2.3	0
K	0.33:1	0.20:1	+0.13	3.4	3.1	+0.3

*Positive sign indicates a higher ratio of initiation to response in the Application group.
†Positive sign indicates that the Application Test generated more discussion.

significant at the two per cent level (two-tailed probability). Also evident were carry-over effects: the clinical teachers tended to hurry through the discussion of whichever test was administered second. However, the balancing of the design ensured that this factor did not affect the overall comparison of the two tests.

DISCUSSION AND CONCLUSIONS

This experiment does not establish directly that feedback from application items produces greater learning than feedback from knowledge items. Nevertheless, in view of the informative role which knowledge of results plays in facilitating learning, the results do suggest a measure of superiority for the higher level items. The lengthier discussion which occurred during the application test tutorials indicates that more information was being exchanged, since the discussion was non-repetitive and relevant to the tested topic. A tentative explanation of this finding is that, since the higher levels in Bloom's Taxonomy subsume the lower ones, the process of 'unpacking' a student's response to a higher level item releases more information than unpacking her response to a lower level one. Thus, despite recent disparagement of the educational relevance of Bloom's cognitive hierarchy (eg, Ormell, 1974), this experiment suggests that position in the Taxonomy is related to a test's usefulness as an instructional aid.

This conclusion is supported by the finding that students in the application test groups had a greater tendency to initiate discussion of new aspects ot the tested topic than those who experienced the knowledge test. This may have been due to the greater relevance of application items to these particular students' interests, for the course investigated is essentially practical. Nevertheless, student initiation of new lines of discussion is likely to facilitate the teacher's task of providing corrective feedback. For it is not likely that a teacher will perceive *all* the weaknesses his students possess, nor that they will *all* be brought to light by test responses alone. If, during discussion of examination results, students (like those in Montor's experiment) seek to defend wrong answers, if they suggest alternative answers to those cited by the teacher, or if they ask questions on topics not mentioned by the teacher, they are in effect lengthening the feedback test—further extending the domain of misconceptions sampled during the feedback process, and thereby increasing the teacher's opportunities to supply information which meets their specific learning needs.

ACKNOWLEDGMENTS—The author gratefully acknowledges the generosity of the Department of Health and Social Security in supporting this work.

REFERENCES

ANGELL, G. W. (1949). The effect of immediate knowledge of quiz results on final examination scores in freshman chemistry. *J. educ. Res.*, 42, 391-394.
ANNETT, J. (1969). *Feedback and Human Behaviour.* Harmondsworth: Penguin.
BEARD, R. (1967). On evaluating the success of teaching. *Br. J. med. Educ.*, 1, 296-302.
BLOOM, B. S., ENGELHART, M. D., FURST, E. J., HILL, W. H., and KRATHWOHL, D. R. (1956). *Taxonomy of Educational Objectives: Handbook 1, Cognitive Domain.* London: Longman.
BOREHAM, N. C. (1977). The use of case histories to assess nurses' ability to solve clinical problems. *J. Adv. Nurs.*, 2, 57-66.
CURTIS, F. D., and WOOD, G. G. (1929). A study of the relative teaching value of four common practices in correcting examination papers. *School Rev.*, 37, 615-623.
ELLIS, J. (1967). Examinations. *Br. J. med. Educ.*, 1, 314-315.
FLANDERS, N. (1970). *Analyzing Teaching Behaviour.* Reading, Mass.: Addison-Wesley.
GILBERT, A. C. F. (1956). Effect of immediacy of knowledge of correctness of response on learning. *J. educ. Psychol.*, 47, 415-423.
MONTOR, K. (1970). Effect of using a self-scoring answer sheet on knowledge retention. *J. educ. Res.*, 63, 435-437.
ORMELL, C. P. (1974). Bloom's Taxonomy and the objectives of education. *Educ. Res.*, 17, 3-18.
PRESSEY, S. L. (1926). A simple apparatus which gives test scores and teaches. *School and Society*, 23, 373-376.
PRESSEY, S. L. (1927). A machine for automatic teaching of drill material. *School and Society*, 25, 549-552.
PRESSEY, S. L. (1950). Development and appraisal of devices providing automatic scoring of objective tests and concomitant self-instruction. *J. Psychol.*, 29, 417-447.
SASSENRATH, J. M., and GARVERICK, C. M. (1965). Effects of differential feedback from examinations on retention and transfer. *J. educ. Psychol.*, 56, 259-263.
SIEGEL, S. (1956). *Nonparametric Statistics for the Behavioural Sciences.* New York: McGraw-Hill.

(Manuscript received April 27, 1977)

AUTHOR INDEX TO VOLUME 47, 1977

ABRAHAMSON, D. S.: See Bell.

ANTHONY, W. S.: Activity in the Learning of Piagetian Operational Thinking 18

ANTHONY, W. S.: The Development of Extraversion and Ability: An Analysis of Rushton's Longitudinal Data 193

BELL, A. E.: Achievement and Self-Reports of Responsibility for Achievement in Informal and Traditional Classrooms 258

BLACKMAN, D. E.: See Newman.

BOLTON, N.: See Zammarelli.

BOREHAM, N. C.: The Effect of Type of Item on Student-Teacher Interaction During Feedback of Examination Performance 335

BROWN, G.: Piagetian Psychology and Education: Time for Revision 7

BUTTERWORTH, G.: See Fairweather.

CAIRNS, E.: The Reliability of the Matching Familiar Figures Test 197

CALNAN, M.: See Ghodsian.

COCHRANE, R.: See Newman.

CONRAD, R.: Lip-Reading by Deaf and Hearing Children 60

CONRAD, R.: The Reading Ability of Deaf School-Leavers 138

COX, M. V.: Teaching Perspective Ability to Five-Year-Olds 312

DESFORGES, C. W.: See Brown.

EISER, C.: Strategies Children Use to Co-ordinate Perspectives as a Function of Task Demands 327

ELLIOTT, C. D.: The Measurement of Speed of Problem Solving and its Relation to Children's Age and Ability 50

ENGELHARDT, J. M.: Analysis of Children's Computational Errors: A Qualitative Approach 149

FAIRWEATHER, H.: The WPPSI at Four Years: A Sex Difference in Verbal-Performance Discrepancies 85

FRANCIS, H.: Children's Strategies in Learning to Read 117

FRANKLIN, B. S.: Effects on Children's Divergent Thinking Abilities of a Period of Direct Teaching for Divergent Production 66

FRANSSON, A.: On Qualitative Differences in Learning IV—Effects of Intrinsic Motivation and Extrinsic Test Anxiety on Process and Outcome 244

GHODSIAN, M.: A Comparative Longitudinal Analysis of Special Education Groups 162

GOLOMB, C.: Symbolic Play: The Role of Substitutions in Pretence and Puzzle Games 175

GOYEN, J. D.: The Relation of Spelling Errors to Cognitive Variables and Word Type 268

GROWSE, R.: See Bell.

HANSFORD, B. C.: An Australian Evaluation of the Junior Eysenck Personality Inventory 330

HARDYCK, C.: Laterality and Intellectual Ability: A Just Not Noticeable Difference? 305

HARGREAVES, D. J.: Sex Roles in Divergent Thinking 25

HARGREAVES, D.: Process of Typification in Classroom Interaction: Models and Methods 274

JARMAN, R. F.: See Lawson.

KEMPA, R. F.: A Study of Attitudes Towards Mathematics in Relation to Selected Student Characteristics.. 296

LAWSON, M. J.: A Note on Jensen's Theory of Level I Ability and Recent Research on Human Memory 91

LAWTON, J. T.: The Development of Causal and Logical Connectives in Children 81

MARTIN, M.: See Goyen.

MASSEY, A. J.: Candidate Fatigue and Performance on G.C.E. Objective Tests 203

McGOUGH, J. M.: See Kempa.

McLAUGHLIN, S. P.: See Wheeler.

MITTLER, P.: See Wheldall.

MOORE, E. M.: Lexical Characteristics of Adolescent Prose: Writer and Task Effects .. 285

MURRAY, D. J.: See Elliott.

NEIDHART, H.: See Hansford.

NEWMAN, C. V.: The Role of Interviews in Choice of University by Psychology Undergraduates.. 209

PUGH, J. G.: See Riding.

RICHARDS, P. N.: See Franklin.

RICHARDSON, K.: Reading Attainment and Family Size: an Anomaly 71

RIDING, R. J.: Iconic Memory and Reading Performance in Nine-Year-Old Children 132

SCOTT, E.: Treatment of Encopresis in a Classroom Setting: A Case Study 199

SEDDON, G. M.: The Effects of Chronological Age on the Relationship of Academic Achievement with Extraversion and Neuroticism: A Follow-up Study 187

SINGLETON, C. H.: Redundancy in the Concept Attainment of Adolescents 40

SMEDSLUND, J.: Piaget's Psychology in Practice 1
SVENSSON, L.: On Qualitative Differences in Learning III—Study Skill and Learning 233
WATKINS, E. J.: See Wheeler.
WATSON, J. M.: The Influence of Context on Referential Description in Children 33
WHEELER, T. J.: Reading Retardation and Cross-Laterality in Relation to Short-Term Information Processing Tasks 126
WHELDALL, K.: On Presenting Pictures and Sentences: The Effect of Presentation Order on Sentence Comprehension in Normal and Mentally Handicapped Children 322
WILSON, J. A.: How Many Teaching Styles? How Much Progress? 105
WONG, B. Y. L.: Further Investigation of the Emergence of Ability for Partial and Complete Multiplication of Classes in Young School Children 76
ZAMMARELLI, J.: The Effects of Play on Mathematical Concept Formation 155

METHUEN

The Psychology and Education of Gifted Children

P. E. VERNON, G. ADAMSON and D. F. VERNON

In these days when education increasingly focuses on average pupils it is all the more important for teachers, parents, and educationists generally to be made aware of the special needs of the bright and talented. The authors of this book discuss the present provision of facilities for the education of these children, with particular reference to the UK and Canada. Detailed consideration is given to the psychological origins of intelligence and creativity, to the available tests and other techniques for identifying exceptionally able children, and to the serious loss that society suffers when the gifted and specially talented are ignored or repressed.

'. . . a lucid, authoritative, thoroughly readable and comprehensive account.'
Times Higher Education Supplement.
£5.95 University Paperback £2.95

The Child with Spina Bifida

ELIZABETH M. ANDERSON and BERNIE SPAIN

The incidence of spina bifida in Britain is one of the highest in the world, and during the last fifteen years the proportion of children surviving has greatly increased. These children require the help and specialised knowledge not only of parents, but of teachers, health visitors, social workers, and physio- and occupational therapists. Drawing on their own and others' research, the authors provide a fully comprehensive and up-to-date account of the range of problems associated with the condition.

'. . . Anderson and Spain's book cannot, I feel sure, fail to become the definitive general text on childhood spina bifida.'
New Statesman
£7.50 Paperback £4.95

Journal of Mental Imagery

Editor: Anees A. Sheikh, Marquette University

This biannual periodical will provide an in-depth analysis of important contemporary imagery topics, drawing on an internationally recognized group of specialists as contributors. Each issue will include experimental, clinical, theoretical and review articles on mental imagery, and will promote interaction between investigators in the field.

Contents of Volume 1, Number 1, Spring 1977

Editor's Introduction, Mental Images: Ghosts of Sensations?
Ahsen, Akhter, Eidetics: An Overview
Bugelski, B. R., Imagery and Verbal Behavior
Cautela, Joseph R., Covert Conditioning: Assumptions and Procedures
Hunter, Ian M. L., Imagery, Comprehension, and Mnemonics
Leuner, Hanscarl, Guided Affective Imagery: An Account of its Development
McKellar, Peter, Autonomy, Mental Imagery, and Dissociation
Richardson, Alan, Verbalizer-Visualizer: A Cognitive Style Dimension
Singer, Jerome L., Imagination and Make-Believe Play in Early Childhood: Some Educational Implications
White, K., Sheehan, P. W., and Ashton, R., Imagery Assessment: A Survey of Self-Report Measures
Yuille, John C., and Catchpole, Michael J., The Role of Imagery in Models of Cognition

Contents of Volume 1, Number 2, Fall 1977

Antrobus, John S., The Dream as a Visual Metaphor Based on Partial Features
Ernest, Carol H., Individual Differences in Imagery Ability: A Critical Review
Lindauer, Martin S., Imagery and Its Place in the Arts, Aesthetics and Creativity
Marks, David, Imagery and Perception: A Theoretical Review
Mowrer, O. Hobart, Mental Imagery: An Indispensable Psychological Construct
Prouty, Garry F., Protosymbolic Method: A Phenomenological Treatment of Schizophrenic Hallucinations
Reese, Hayne W.. Toward a Cognitive Theory of Mnemonic Imagery
Reyher, Joseph, Spontaneous Visual Imagery: Implications for Psychoanalysis, Psychopathology and Psychotherapy
Sarbin, Theodore R., Social Context of Hallucinations
Sheikh, Anees A., and Richardson, P. M., Mental Imagery and Psychosomatic Illness: A Critical Review
Tomkins, Silvan S., The Central Role of Imagery in Perception, Cognition and Action
Virel, Andre, Genesis of an Image
Zuckerman, Marvin, and Tushup, R., The Effects of Stimulus Invariance on Daydreaming and Divergent Thinking

Editorial Board: Akhter Ahsen, John S. Antrobus, Rudolf Arnheim, Fred Attneave, Theodore X. Barber, Aaron T. Beck, B. R. Bugelski, Joseph R. Cautela, Leonard W. Doob, Carol H. Ernest, E. T. Gendlin, Ralph N. Haber, Ian Hunter, Charles S. Jordan, Eric Klinger, F. Paul Kosbab, Stephen Kosslyn, Hanscarl Leuner, Martin S. Lindauer, David Marks, Peter McKellar, O. Hobart Mowrer, Garry F. Prouty, Hayne W. Reese, Joseph Reyher, Alan Richardson, Theodore R. Sarbin, Sohan L. Sharma, Peter W. Sheehan, Roger N. Shepard, Jerome L. Singer, Silvan S. Tomkins, Andre Virel, A. Daniel Yarmey, John C. Yuille, Marvin Zuckerman.

Brandon House, Inc., Orders: P.O. Box 240, Bronx, New York 10471

☐ Individual Subscription—Check enclosed	$15.00
☐ Institutional Subscription	$20.00

phlo 7

name ____________________
address ____________________
city ____________________ state __________ zip __________

NEW VOLUMES IN THE
EDUCATIONAL PSYCHOLOGY
SERIES

PLAYFULNESS

Its relationship to Imagination and Creativity

J. NINA LIEBERMAN

May 1977, 192pp, £8.50/$12.00 0.12.449450.1

Playfulness — how we play as opposed to why we play — is a clue to cognitive style, and goes beyond the childhood years to exist as a personality trait in adolescents and adults. *Playfulness: Its Relationship to Imagination and Creativity* develops the concept of playfulness, tests instruments to identify and measure it, and explores the connections between playfulness, as prosocial behaviour, and imagination and creativity at all age levels.

The book is perhaps the first to deal comprehensively with the question of how we play as an indication of our behaviour in general. The author presents a theoretical model of playfulness including research based on the model. Exploratory behaviour and play are differentiated, and humour, joy and spontaneity are isolated as the three components of playfulness. The role of playfulness in everyday living, and its relations to childrearing practices, educational planning, career choice and job satisfaction, and the use of leisure are all traced. The practical implications of this theory and research for educational and developmental psychologists, and educators at all levels are discussed and a basis for future research supplied.

TELEVISION AND THE PRESCHOOL CHILD

A Psychological Theory of Instruction and Curriculum Development

HARVEY LESSER

September 1977, 288pp, £11.35/$16.00 0.12.444250.1

Television and the Preschool Child documents the surprising claim that television is not as powerful a social and educational force as is commonly believed. The book refutes the widely accepted research findings that a relationship exists between viewing violence and behaving violently and questions the research that claims *Sesame Street* is effective, especially for its target population of disadvantaged preschool children.

The central tenet of the book is that the educational potential of television for preschool children is largely untapped. *Television and the Preschool Child* provides the clearest and most detailed exposition to date of the problems of instructing young children through television. American, Soviet, and Swiss (Piagetian) research on the development of cognition, language and memory in children, is reviewed, presenting a proposal, grounded on the research, for improving television instruction for preschool infants.

The book also features a general introduction to the history and economics of children's television, discussion of the limits imposed on instruction by the passive qualities of viewing, and an examination of such new developments as computer-assisted instruction in their relation to television.

Academic Press London New York San Francisco
A Subsidiary of Harcourt Brace Jovanovich, Publishers
24-28 Oval Road, London NW1, England
111 Fifth Avenue, New York, NY 10003, USA

NEW VOLUMES IN THE
EDUCATIONAL PSYCHOLOGY
SERIES

PSYCHOLOGICAL PROCESSES IN EARLY EDUCATION

edited by HARRY L. HOM, Jr. and PAUL A. ROBINSON

July 1977, 352pp, £12.40/$17.50 0.12.354450.5

To extend contemporary psychological research and theory to early childhood education is the goal of this treatise consisting of original chapters by distinguished child psychologists and educators. For each topic, the contributors review research data and theory and discuss the implications for early childhood education. The book offers perhaps the first comprehensive coverage addressed specifically to early childhood education of prosocial behaviours, punishment, language interaction between teacher and child, individual differences in cognitive style, and behaviour disorders in the preschool-age child.

The papers are divided into three sections. The first section gives an overview of early education and its relation to child psychology. The second section deals with basic socialisation processes in child development — including modelling, punishment, behaviour analysis, and incentives — and their application in early childhood settings. The third section features papers covering fundamental aspects of child behaviour and development such as cognitive development, prosocial behaviour, language, and behaviour disorders.

HANDBOOK ON TEACHING EDUCATIONAL PSYCHOLOGY

edited by DONALD J. TREFFINGER, J. KENT DAVIS and RICHARD E. RIPPLE

August 1977, 384pp, £12.80/$18.00 0.12.697750.X

This book is perhaps the first to address the growing recognition among educational psychologists that they are charged not only with studying and improving the teaching and learning of *others,* but with looking as well at the quality of *their own* instruction. It provides a contemporary, wide-ranging survey of practices and problems in teaching educational psychology and reviews and evaluates the conceptual and methodological bases of practices and problems. As a survey and review of contemporary educational psychology *as it is taught,* the book brings together in a single reference source-book many themes previously treated only in an isolated and fragmented way.

The book's four sections each contain chapters prepared by leaders in various aspects of educational psychology. The diversity of approaches in the various chapters ensures an accurate representation of the field. Major issues and advances in teaching educational psychology at both graduate and undergraduate level are reviewed, together with specific teaching approaches and strategies currently employed. The general problems confronting educational psychologists and the methods of dealing with them are investigated. At the same time the influences on research and instructional development in this field are discussed.

Academic Press London New York San Francisco
A Subsidiary of Harcourt Brace Jovanovich, Publishers
24-28 Oval Road, London NW1, England
111 Fifth Avenue, New York, NY10003, USA

PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

available from the Society's offices

St. Andrews House, 48 Princess Road East, Leicester, LE1 7DR

Careers in Psychology	50p
Psychology and Psychologists: A Pamphlet for Schools	15p
Teaching Educational Psychology in Training Colleges	25p
Technical Recommendations for Psychological and Educational Tests	10p
Aspects of Autism. Edited by P. J. Mittler	50p
The Psychological Study of Deafness and Hearing Impairment. Edited by M. Rodda	40p
Postgraduate Qualifications and Courses in Psychology 1977/8	40p

PRESIDENTIAL ADDRESSES

1968 Dissonance and Stress in the Teaching of Psychology: Some Reflections of an Accidental Psychologist: Professor George Seth	10p
1969 Changing Horizons: An Essay in Autobiography: Dr. Boris Semeonoff	10p
1970 Choosing: Professor R. J. Audley	10p
1971 In Search of an Idiographic Psychology. By Professor H. Gwynne Jones	10p
1972 Psychology today and tomorrow: Professor Harry Kay	10p
1973 Psychology in Society: Ends or End: Professor Max Hamilton	10p
1974 On Taking Sides: Professor B. M. Foss	10p
1975 Thought and the Brain: Professor O. L. Zangwill	10p
1976 Psychology and Social Policy: Professor Jack Tizard	10p
1977 The scientific/applied debate in psychology: A contribution. May A. Davidson	20p

C. S. MYERS LECTURES

Remembering Dr. Myers: Sir Frederic Bartlett	10p
Beyond Industrial Psychology: Dr. R. Conrad	10p
Practice at Liking: A Study in Experimental Aesthetics: Professor D. W. Harding	10p
C. S. Myers in Retrospect: Professor Alec Rodger	10p
On Applied Psychology in Britain: Professor G. C. Drew	10p
Measuring Psychiatric Morbidity in a South Indian Population: Professor G. M. Carstairs	10p

MEMORANDA OF EVIDENCE

To the Royal Commission on the Law Relating to Mental Illness and Mental Deficiency 1954-1957	10p
To the Royal Commission on the Penal System in England and Wales	10p
To the Royal Commission on Medical Education	10p
To the Ministry of Health and Department of Education and Science on the Responsibility for the Education of Mentally Handicapped Children	10p
To the Advisory Council on the Penal System (Sub-committee on Detention Centres)	10p
To the Departmental Committee on the Adoption of Children	10p
To the Department of Health and Social Security Sub-committee on the Role of Psychologists in the Health Services	10p
To the Butler Committee on the Law relating to the Mentally Abnormal Offender	10p
To the Warnock Committee on the Enquiry into Special Education (Summary)	10p
To the Royal Commission on the N.H.S	35p
To the Data Protection Committee (Home Office)	15p

MISCELLANEOUS

Principles governing the employment of psychological tests and clinical instruments	7p
Psychological Tests: A Statement by the British Psychological Society	7p
Memorandum on the Establishment and Accommodation of Psychologists in General Hospitals	7p
The School Psychological Service (England and Wales)	7p
The changing role of the Educational Psychologist in Child Guidance Services in Scotland	10p
Teaching Psychology to Nurses (DCP Report)	10p
Training in Psychotherapy (DCP Report)	10p
Guidelines for the In-service Training of Probationer Clinical Psychologists	10p

(All prices on the above publications include postage)

The British Journal of Medical Psychology

Volume 50, Part I, March 1977

Ruth Purisman and Benjamin Maoz. Adjustment and war bereavement—Some considerations
Heinz H. Wolff. Loss: A central theme in psychotherapy
Miller Mair. Psychology and psychotherapy: Some common concerns
H. P. Hildebrand. Psychotherapy by non-medical psychotherapists: A non-medical point-of-view
W. H. Trethowan. A psychiatrist's view of lay psychotherapy
T. Trauer. Attitudes of psychiatric patients to staff roles and treatment methods: A replication and extension
Geoff Shepherd and Robert Durham. The multiple techniques approach to behavioural psychotherapy: A retrospective evaluation of effectiveness and an examination of prognostic indicators
Freda E. Martin. Some implications from the theory and practice of family therapy for individual therapy (and *vice versa*)
Jonathan C. Sperber. Personal constructs and child psychiatric diagnosis—A pilot study
G. A. Foulds and A. Bedford. Hierarchies of personality deviance and personal illness
Robert M. Tutone. Correlates of illness susceptibility
Nils Cochrane. A contribution from the field of psychodynamics to the controversy over reactive and endogenous depression
Mualla Ozturk and Orhan Ozturk. Thumbsucking and falling asleep
Armando R. Favazza. Feral and isolated children

Volume 50 (1977) £16·00 ($39·50 USA and Canada). Single parts £5·00 ($11·75 USA and Canada) each.
Enquiries about advertising in the journal should be sent to the London or New York office of the Press.

CAMBRIDGE UNIVERSITY PRESS
BENTLEY HOUSE, 200 EUSTON ROAD, LONDON, NW1 2DB
32 EAST 57TH STREET, NEW YORK, N.Y. 10022

The British Journal of Psychology

Volume 68, Part 3, August 1977

M. O'MAHONEY and H. MUHIUDEEN. A preliminary study of alternative taste languages using qualitative description of sodium chloride solutions: Malay versus English

DONALD P. SPENCE and GUDMUND J. W. SMITH. Experimenter bias against subliminal perception? Comments on a replication

PAUL J. BARBER. Experimenter bias against subliminal perception? A rejoinder

GÖRAN H. BERGER and SAMUEL C. B. GAUNITZ. Self-rated imagery and vividness of task pictures in relation to visual memory

JON WRIGHT, DONALD S. CICCONE and JOHN W. BRELSFORD. Selective encoding processes in recognition

JEFFREY FIELD and JOHN K. COLLINS. A developmental study of the effects of instructions on visual shape judgements

ROBERT T. SOLMAN. Evidence that focal processing involves a build-up of a visual object

BO S. JOHANSSON. Levels of mastery of the coordinators *and* and *or* and logical test performance

ROBERT WOOD and DOUGLAS T. WILSON. A technique for converting ranks into measures

J. R. NICHOLSON and G. M. SEDDON. The influence of secondary depth cues on the understanding by Nigerian schoolboys of spatial relationships in pictures

SAFAR DAEE and J. M. WILDING. Effects of high intensity white noise on short-term memory for position in a list and sequence

PETER E. MORRIS. A criticism of Wilkins' (1971) measure of category size, and its implication for the Smith, Shoben & Rips (1974) model of semantic memory

SUSAN AYLWIN. The structure of visual and kinaesthetic imagery: A free association study

D. W. J. CORCORAN and T. G. HOUSTON. Is the lemon test an index of arousal level?

J. E. TONG, GILLIAN LEIGH, J. CAMPBELL and D. SMITH. Tobacco smoking, personality and sex factors in auditory viligance performance

JOHN HINTON. A psychophysiological study of paranoid hostility and defensiveness in maximum security hospital patients

SUSANNA MILLAR. Tactual and name matching by blind children

BOOK REVIEWS

OTHER PUBLICATIONS RECEIVED

Volume 68 (1977) £20.00 ($49.50 USA and Canada). Single parts £6.00 ($14.00 USA and Canada) each

Enquiries about advertising in the Journal should be sent to the London or New York office of the Press

CAMBRIDGE UNIVERSITY PRESS

Bentley House, 200 Euston Road, London NW1 2DB
32 East 57th Street, New York, N.Y. 10022

The British Journal of Medical Psychology

Volume 50, Part 2, June 1977

Douglas Hooper, F. John Roberts, Mary K. Hinchcliffe and Pamela W. Vaughan. The melancholy marriage: An inquiry into the interaction of depression I. Introduction

Mary K. Hinchcliffe, Pamela W. Vaughan, Douglas Hooper and F. John Roberts. The melancholy marriage: An inquiry into the interaction of depression II. Expressiveness

W. S. Inman. Sigmund Freud, FRS meets Isaac Newton, FRS

Phyllis Shaw. A study of social problems in a group of young women treated with brief Psychotherapy

A. Bedford and G. A. Foulds. Validation of the Delusions-Symptoms-States Inventory

J. Adams-Webber. The organization of judgements based on positive and negative adjectives in the Bannister-Fransella Grid Test

K. Bragan. The psychological gains and losses of religious conversion

John Schwartzman. Systemic aspects of abstinence and addiction

Per-Ake Magnusson, Alf Nilsson and Nils-Gunnar Henriksson. Psychogenic vertigo within an anxiety frame of reference: An experimental study

Review: Eysenck Personality Questionnaire (Junior and Adult)

Books received

Volume 50 (1977) £16.00 ($39.50 USA and Canada). Single parts £5.00 ($11.75 USA and Canada) each.

Enquiries about advertising in the Journal should be sent to the London or New York office of the Press.

CAMBRIDGE UNIVERSITY PRESS

BENTLEY HOUSE, 200 EUSTON ROAD, LONDON NW1 2DB
32 EAST 57TH STREET, NEW YORK, N.Y. 10022

EDUCATIONAL RESEARCH

Vol. 20, No. 1 *Editor* : ALFRED YATES *November*, 1977

JOHN RAVEN. School rejection and its amelioration.

A. MACMILLAN and I. KOLVIN. Behaviour modification in teaching strategy: some emergent problems and suggested solutions.

JANE ROOT. The importance of peer groups.

ROY EDGLEY. Education for industry.

IAN MACLENNAN An hypothesis of thinking.

DELROY M. LOUDEN. Conflict and change among West Indian parents and their adolescents in Britain.

Debate on the place of inferential statistics in educational research:

P. F. W. PREECE. A note in defence of inferential statistics.

M. B. YOUNGMAN. Necessary inferences: a reply to T. Derrick.

GABRIEL CHANAN. Statistics debate masks analytical weakness.

PETER MORTIMORE. Schools as institutions.

BOOK REVIEWS.

The subscription is £5·25 a year (incl. postage); £7·25 Overseas (postage paid).

All orders should be sent to :

NFER Publishing Company Ltd. (Journals Department), Darville House,
2 Oxford Road East, Windsor, Berks, SL4 1QS

The British Journal of Psychology

Volume 68, Part 1, February, 1977

STEPHEN FINEMAN. The achievement motive construct and its measurement : Where are we now?
L. C. MORRISON. Inappropriate constancy scaling as a factor in the Muller-Lyer illusion
ALAN RICHARDSON. The meaning and measurement of memory imagery
SIMON FOLKARD, TIMOTHY H. MONK, ROSAMUND BRADBURY and JOANNA ROSENTHALL. Time of day effects in school children's immediate and delayed recall of meaningful material
H. B. GIBSON, M. E. CORCORAN and J. D. CURRAN. Hypnotic susceptibility and personality : The consequences of a diazepam and the sex of the subjects
ELIZABETH M. ANDERSON and IAN PLEWIS. Impairment of a motor skill in children with spina bifida cystica and hydrocephalus : An exploratory study
N. E. WETHERICK and JANE ALEXANDER. The role of semantic information in short-term memory in children aged 5 to 9 years
JEAN-LUC LAMBERT. Sequential contrast effects with retarded subjects after discrimination learning with and without errors
MARGERY HEBER. The influence of language training on seriation of 5-6 year old children initially at different levels of descriptive competence
J. G. WOLFF. The discovery of segments in natural language
J. COSTERMANS and M. HUPET. The other side of Johnson-Laird's interpretation of the passive voice
P. N. JOHNSON-LAIRD. The Passive Paradox : A reply to Costermans and Hupet
JOHN A. SLOBODA. Phrase units as determinants of visual processing in music reading
JOHN ARCHER. Sex differences in the emotional behaviour of laboratory mice
ANGELA SUMMERFIELD and J. A. LAKE. Non-verbal and verbal behaviours associated with parting
BOOK REVIEWS
OTHER PUBLICATIONS RECEIVED

Volume 68 (1977) £20·00 ($49·50 USA and Canada). Single parts £6·00 ($14·00 USA and Canada) each
Enquiries about advertising in the Journal should be sent to the London or New York office of the Press

CAMBRIDGE UNIVERSITY PRESS

Bentley House, 200 Euston Road, London, NW1 2DB.
32 East 57th Street, New York, N.Y. 10022

THE BRITISH JOURNAL OF SOCIAL AND CLINICAL PSYCHOLOGY

Volume 16, Part 1, February, 1977

MEIR TEICHMAN. Affiliative behaviours among soldiers during war-time
JOHN M. INNES. Extremity and 'don't know' sets in questionnaire response
JAMES M. OLSON and JOHN T. PARTINGTON. An integrative analysis of two cognitive models of interpersonal effectiveness
GORDON GRAIG and STEVEN W. DUCK. Similarity, interpersonal attitudes and attraction: The evaluative-descriptive distinction
M. GHODSIAN. Children's behaviour and the BSAG: Some theoretical and statistical considerations
P. E. BULL and R. BROWN. The role of postural change in dyadic conversations
PETER WARR and PAUL JACKSON. Salience, importance and evaluation in judgments about people.
D. R. RUTTER. Speech patterning in recently admitted and chronic long-stay schizophrenic patients
S. B. G. EYSENCK and H. J. EYSENCK. The place of impulsiveness in a dimensional system of personality description
J. VILA and H. R. BEECH. Vulnerability and conditioning in relation to the human menstrual cycle
EDGAR MILLER. The management of dementia: A review of some possibilities
R. P. LIBERMAN, C. H. FEARN, W. DERISI, J. ROBERTS and M. CARMONA. The Credit-Incentive System: Motivating the participation of patients in a day hospital
JOHN C. TOUHEY. Machiavellianism and teachers' perceptions of elementary school pupils
M. J. KIRTON. Ray's balanced Dogmatism scale re-examined
EDGAR MILLER. A note on visual information processing in presenile dementia: A preliminary report
BOOK REVIEWS
PUBLICATIONS RECEIVED

Volume 16 (1977) £19·00 ($47·50 *USA and Canada*). *Single parts* £6·00 ($14·00 *USA and Canada*) *each.*

Enquiries about advertising in the Journal should be sent to the London or New York office of the Press.

CAMBRIDGE UNIVERSITY PRESS

BENTLEY HOUSE, 200 EUSTON ROAD, LONDON, NW1 2DB
32 EAST 57th STREET, NEW YORK, N.Y. 10022

THE BRITISH JOURNAL OF SOCIAL AND CLINICAL PSYCHOLOGY

Volume 16, Part 3, September 1977

K. J. HESKIN, N. BOLTON, P. A. BANISTER and F. V. SMITH. Prisoners' personality: A factor analytically derived structure
J. K. DUA. Effect of audience on the acquisition and extinction of avoidance
ROBERT B. BURNS. Male and female perceptions of their own and the other sex
CLARE A. HYMAN. A report on the psychological test results of battering parents
GRAHAM F. WAGSTAFF. An experimental study of compliance and post-hypnotic amnesia
STEVE DUCK and GORDON CRAIG. The relative attractiveness of different types of information about another person
MARIE JOHNSTON and ANNE HACKMANN. Cross-validation and response sets in repeated use of mood questionnaires
ERIK WOODY and GORDON CLARIDGE. Psychoticism and thinking
J. P. N. PHILLIPS. On the incorrect investigation of interactions
A. W. HEIM, S. M. UNWIN and K. P. WATTS. An investigation into disordered adolescents by means of the Brook Reaction Test
D. G. BYRNE, DESLEY BOYLE and D. W. PRITCHARD. Sex differences in response to a self-rating depression scale
ANTONIA WHITEHEAD and ANDREW MATHEWS. Attitude change during behavioural treatment of sexual inadequacy
GEOFFREY BEATTIE. The dynamics of interruption and the filled pause
D. HOWITT, G. CRAVEN, CAROL IVESON, J. McCABE and T. ROLPH. The misdirected letter
S. H. A. HENLEY, N. F. DIXON and A. E. CARTMELL. A note on the relationship between authoritarianism and acceptance of military ideology
BOOK REVIEWS
PUBLICATIONS RECEIVED

Volume 16 (1977) £19.00 ($47.50 USA and Canada). Single parts £6.00 ($14.00 USA and Canada) each.

Enquiries about advertising in the Journal should be sent to the London or New York office of the Press.

CAMBRIDGE UNIVERSITY PRESS

BENTLEY HOUSE, 200 EUSTON ROAD, LONDON NW1 2DB
32 EAST 57th STREET, NEW YORK, N.Y. 10022

PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

The British Journal of Psychology
The British Journal of Medical Psychology
Journal of Occupational Psychology
The British Journal of Social and Clinical Psychology
The British Journal of Mathematical and Statistical Psychology
The British Journal of Educational Psychology

The British Journal of Psychology, *The British Journal of Medical Psychology*, *Journal of Occupational Psychology* and *The British Journal of Social and Clinical Psychology* appear in quarterly parts. The 1977 subscription prices, payable in advance, are £20·00 (US $60·00 in the USA and Canada) for *The British Journal of Psychology* ; £19·00 (US $57·00) for *The British Journal of Social and Clinical Psychology*; £16·00 (US $48·00) for *The British Journal of Medical Psychology* and £13·00 (US $39·00) for *Journal of Occupational Psychology*. Subscription to any of these journals may be sent to any bookseller or subscription agent, or to the Cambridge University Press, Bentley House, 200 Euston Road, London, NW1 2DB. American and Canadian enquiries should be addressed to the Cambridge University Press, American Branch, 32 East 57th Street, New York, N.Y.10022.

The British Journal of Mathematical and Statistical Psychology is issued twice a year, in May and November. The subscription rate, payable in advance, is £15·00 per volume post free (US $34·00 in the USA). Subscriptions should be sent to the Society's agents, John Wright and Sons, Ltd., 42-44 Triangle West, Bristol, BS8 1EX, or to any bookseller.

The British Journal of Educational Psychology is issued three times a year, in February, June and November. The prepaid subscription is £8 a year, post free, or £3·00 each number. The prepaid subscription price in the USA is $20·00 post free. Subscriptions should be sent to Scottish Academic Press, 33 Montgomery Street, Edinburgh, EH7 5JX.

The series of *Monograph Supplements* issued in connection with *The British Journal of Psychology* is now discontinued. Orders for back volumes may be sent to any bookseller, or the Cambridge University Press, Bentley House, 200 Euston Road, London, NW1, from whom details of these publications may be obtained.

MANUSCRIPTS

For *The British Journal of Psychology* should be sent to Professor A. D. B. Clarke, Department of Psychology, The University, Hull, HU6 7RX;

For *The British Journal of Medical Psychology* should be sent to Professor Arthur H. Crisp, Psychiatric Research Unit, Atkinson Morley's Hospital, Copse Hill, London, SW20 0NE;

For *Journal of Occupational Psychology* should be sent to Professor B. Shackel, Department of Human Sciences, University of Technology, Loughborough, Leicestershire;

For *The British Journal of Social and Clinical Psychology* should be sent to Dr. Halla Beloff (Social Psychology Editor), Department of Psychology, University of Edinburgh, 60 Pleasance, Edinburgh, EH8 9TJ; or to Dr. H. R. Beech (Clinical Psychology Editor), Department of Psychiatry, Withington Hospital, West Didsbury, Manchester, M20 8LR;

For *The British Journal of Mathematical and Statistical Psychology* should be sent to Professor P. Levy, Department of Psychology, University of Lancaster, Lancaster, LA1 4YF;

For *The British Journal of Educational Psychology* should be sent to Professor N. J. Entwistle, Department of Educational Research, University of Lancaster, Lancaster.

Contributors are advised to consult *Suggestions to Authors* issued by the Standing Committee on Publications of the British Psychological Society and obtainable, price £0·50 (USA $2·00) post free, from Cambridge University Press.

All manuscripts should be submitted in duplicate.

OFF-PRINTS

Contributors on application receive fifty copies of their papers free. Additional copies may be obtained in units of fifty copies.

SUBSCRIPTIONS TO THE BRITISH PSYCHOLOGICAL SOCIETY

The subscription for Fellows, Associates and Graduate Members is £18·00 (£16·00 if paid by Banker's Order).

They are entitled to purchase any or all of the Society's Journals at the special rate of £4·00 per Journal per annum.

Information on membership and activities may be obtained from the Secretary, The British Psychological Society, St. Andrew's House, 48 Princess Road, Leicester, LE1 7DQ.

CONTENTS

SYMPOSIUM: LEARNING PROCESSES AND STRATEGIES—III.

SVENSSON, L.:	On Qualitative Differences in Learning: III—Study Skill and Learning	233
FRANSSON, A.:	On Qualitative Differences in Learning: IV—Effects of Intrinsic Motivation and Extrinsic Test Anxiety on Process and Outcome	244

1 6 JUN 1978

ARTICLES:

BELL, A. E., ABRAHAMSON, D. S., and GROWSE, R.	Achievement and Self-Reports of Responsibility for Achievement in Informal and Traditional Classrooms	258
GOYEN, J. D., and MARTIN, M.:	The Relation of Spelling Errors to Cognitive Variables and Word Type	268
HARGREAVES, D.:	The Process of Typification in Classroom Interaction: Models and Methods	274
MOORE, E. M.:	Lexical Characteristics of Adolescent Prose: Writer and Task Effects	285
KEMPA, R. F., and McGOUGH, J. M.:	A Study of Attitudes Towards Mathematics in Relation to Selected Student Characteristics	296
HARDYCK, C.:	Laterality and Intellectual Ability: A Just Not Noticeable Difference?	305
COX, M. V.:	Teaching Perspective Ability to Five-Year-Olds	312

RESEARCH NOTES:

WHELDALL, K., and MITTLER, P.	On Presenting Pictures and Sentences: The Effect of Presentation Order on Sentence Comprehension in Normal and Mentally Handicapped Children	322
EISER, C.:	Strategies Children Use to Co-ordinate Perspectives as a Function of Task Demands	327
HANSFORD, B. C., and NEIDHART, H.	An Australian Evaluation of the Junior Eysenck Personality Inventory	330
BOREHAM, N. C.:	The Effect of Type of Item on Student-Teacher Interaction During Feedback of Examination Performance	335

AUTHOR INDEX TO VOLUME 47 339